



The Routledge Handbook of Instructed Second Language Acquisition

Edited by Shawn Loewen and Masatoshi Sato

The Routledge Handbook of Instructed Second Language Acquisition

The Routledge Handbook of Instructed Second Language Acquisition is the first collection of state-of-the-art papers pertaining to Instructed Second Language Acquisition (ISLA). Written by 45 world-renowned experts, the entries are full-length articles detailing pertinent issues with up-to-date references. Each chapter serves three purposes:

- (1) provide a review of current literature and discussions of cutting edge issues;
- (2) share the authors' understanding of, and approaches to, the issues; and
- (3) provide direct links between research and practice.

In short, based on the chapters in this handbook, ISLA has attained a level of theoretical and methodological maturity that provides a solid foundation for future empirical and pedagogical discovery. This handbook is the ideal resource for researchers, graduate students, upper-level undergraduate students, teachers, and teacher-educators who are interested in second language learning and teaching.

Shawn Loewen is Associate Professor of Second Language Studies at Michigan State University, USA.

Masatoshi Sato is Associate Professor of Applied Linguistics at Universidad Andrés Bello, Chile.

Routledge Handbooks in Applied Linguistics

Routledge Handbooks in Applied Linguistics provide comprehensive overviews of the key topics in applied linguistics. All entries for the handbooks are specially commissioned and written by leading scholars in the field. Clear, accessible and carefully edited *Routledge Handbooks in Applied Linguistics* are the ideal resource for both advanced undergraduates and postgraduate students.

For a full list of titles in this series, please visit www.routledge.com/Routledge-Handbooks-in-Applied-Linguistics/book-series/RHAL

The Routledge Handbook of World Englishes

Edited by Andy Kirkpatrick

The Routledge Handbook of Applied Linguistics

Edited by James Simpson

The Routledge Handbook of Discourse Analysis

Edited by James Paul Gee and Michael Handford

The Routledge Handbook of Second Language Acquisition

Edited by Susan Gass and Alison Mackey

The Routledge Handbook of Language and Intercultural Communication

Edited by Jane Jackson

The Routledge Handbook of Language Testing

Edited by Glenn Fulcher and Fred Davidson

The Routledge Handbook of Multilingualism

Edited by Marilyn Martin-Jones, Adrian Blackledge and Angela Creese

The Routledge Handbook of Translation Studies

Edited by Carmen Millán-Varela and Francesca Bartrina

The Routledge Handbook of Language and Health Communication

Edited by Heidi E. Hamilton and Wen-ying Sylvia Chou

The Routledge Handbook of Language and Professional Communication

Edited by Stephen Bremner and Vijay Bhatia

The Routledge Handbook of Instructed Second Language Acquisition

*Edited by
Shawn Loewen and Masatoshi Sato*

 **Routledge**
Taylor & Francis Group
NEW YORK AND LONDON

First published 2017
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2017 Taylor & Francis

The right of Shawn Loewen and Masatoshi Sato to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-93623-2 (hbk)
ISBN: 978-1-315-67696-8 (ebk)

Typeset in Times New Roman
by Apex CoVantage, LLC

Contents

<i>List of Illustrations</i>	viii
<i>List of Contributors</i>	x
 1 Instructed Second Language Acquisition (ISLA): An Overview <i>Shawn Loewen and Masatoshi Sato</i>	 1
SECTION I	
Second Language Processes and Products	13
2 Knowledge and Skill in ISLA <i>Robert DeKeyser</i>	15
3 Intentional and Incidental L2 Learning <i>Ronald P. Leow and Celia C. Zamora</i>	33
4 Complexity, Accuracy, and Fluency in L2 Production <i>Marije Michel</i>	50
5 Sociocultural Theory in the L2 Classroom <i>Neomy Storch</i>	69
SECTION II	
Approaches to Second Language Instruction	85
6 Content-Based Language Teaching <i>Roy Lyster</i>	87
7 Task-Based Language Teaching <i>Rod Ellis</i>	108
8 Cognitive-Interactionist Approaches to L2 Instruction <i>YouJin Kim</i>	126

Contents

9	Concept-Based Language Instruction <i>James P. Lantolf and Xian Zhang</i>	146
10	Processing Instruction <i>Bill VanPatten</i>	166
11	Assessment in the L2 Classroom <i>Ute Knoch and Susy Macqueen</i>	181
SECTION III		
Language and Instructed Second Language Acquisition		203
12	Grammar Acquisition <i>Hossein Nassaji</i>	205
13	Acquisition of L2 Pragmatics <i>Kathleen Bardovi-Harlig</i>	224
14	L2 Fluency Development <i>Tracey M. Derwing</i>	246
15	Pronunciation Acquisition <i>Sara Kennedy and Pavel Trofimovich</i>	260
16	Vocabulary Acquisition <i>Beatriz González-Fernández and Norbert Schmitt</i>	280
17	Written Language Learning <i>Charlene Polio and Jongbong Lee</i>	299
SECTION IV		
Instructed Second Language Acquisition Learning Environments		319
18	ISLA in East Asian Contexts <i>Yuko Goto Butler</i>	321
19	Study Abroad and ISLA <i>Carmen Pérez-Vidal</i>	339
20	Computer-Assisted SLA <i>Hayo Reinders and Glenn Stockwell</i>	361

SECTION V	
Individual Differences and Instructed Second Language Acquisition	377
21 Social Dimensions and Differences in Instructed SLA <i>Patricia A. Duff</i>	379
22 Cognitive Differences and ISLA <i>Shaofeng Li</i>	396
23 Motivation in the L2 Classroom <i>Kata Csizér</i>	418
24 Psychological Dimensions and Foreign Language Anxiety <i>Jean-Marc Dewaele</i>	433
25 L2 Instructor Individual Characteristics <i>Laura Gurzynski-Weiss</i>	451
26 Child ISLA <i>Rhonda Oliver, Bich Nguyen, and Masatoshi Sato</i>	468
27 Instructed Heritage Language Acquisition <i>Silvina Montrul and Melissa Bowles</i>	488
SECTION VI	
Instructed Second Language Acquisition Research Methods	503
28 Quantitative Research Methods <i>Luke Plonsky</i>	505
29 Qualitative Research Methods <i>Peter I. De Costa, Lorena Valmori, and Ina Choi</i>	522
30 Classroom-Based Research <i>Alison Mackey</i>	541
31 Experimental Research Methods <i>Kim McDonough</i>	562
32 Ethics in ISLA <i>Susan Gass and Scott Sterling</i>	577
<i>Index</i>	597

Illustrations

Figures

6.1	Range of CBLT settings	88
6.2	Instructional sequence integrating language and content in CBLT	99
9.1	Partial SCOPA for particle “out,” meaning “selection”	153
9.2	SCOPA for topicalization of object	157
9.3	Material SCOPA illustrating topicalization options in Chinese	157
9.4	Kris’s gesture while producing an OSV utterance	160
24.1	Nested design of anxieties	439
25.1	The interplay of contextual factors, learner individual differences, and instructor characteristics in ISLA	453
25.2	Developing a taxonomy of instructor corrective feedback decision-making	461
28.1	Basic design scheme and major decisions in ISLA research	506
28.2	Main contrasts in ISLA analyses	510
28.3	Summary of meta-analytic effects in ISLA across subdomains	515
29.1	Multifaceted nature of language learning and teaching	523
29.2	Aligning methodology with theory and paradigm	526

Tables

7.1	Comparison of task-supported and task-based language teaching	110
7.2	Differences in task-based language teaching approaches	112
7.3	Addressing some common misconceptions about TBLT	114
7.4	Typical design variables	119
11.1	Dimensions of CBA with examples	183
11.2	Teacher considerations relating to different assessment methods based on Hill and McNamara’s (2012) dimensions of CBA	195
14.1	Common measures of utterance fluency	247
16.1	Nation’s (2013) framework of the dimensions involved in knowing a word	283
20.1	Affordances of CALL	363
23.1	Classroom-related elements from Dörnyei and Ottó’s (1998) process model of L2 motivation	421
28.1	Three critiques of <i>p</i> -values and corresponding benefits of effect sizes	510

Illustrations

28.2	Descriptive statistics for sample study	511
28.3	Descriptors for meta-analyses with multiple groups presented in Figure 28.3	516
29.1	Paradigms, epistemology, and ontology	526
32.1	RCR domains	581

Contributors

Kathleen Bardovi-Harlig is Professor of Second Language Studies at Indiana University where she teaches and conducts research on second language acquisition, L2 pragmatics, and tense-aspect systems. Her work on pragmatics has appeared in *Language Learning*, *SSLA*, *Intercultural Pragmatics*, and *Language Teaching Research*. She is co-editor of *Interlanguage Pragmatics* and *Teaching Pragmatics* and co-editor of two volumes of *Pragmatics and Language Learning* (2006, 2016).

Melissa Bowles is Associate Professor of Spanish, Linguistics, and Educational Psychology and Director of the Second Language Acquisition and Teacher Education (SLATE) PhD concentration at the University of Illinois at Urbana-Champaign. Her main research interests are classroom second and heritage language acquisition and the ways in which instruction differentially affects the two learner groups.

Yuko Goto Butler is Associate Professor of Educational Linguistics at the Graduate School of Education at the University of Pennsylvania. She is also the director of Teaching English to Speakers of Other Languages (TESOL) program at Penn.

Ina Choi is a PhD candidate in Second Language Studies at Michigan State University. She holds an MA in Foreign and Second Language Education from The Ohio State University. Her research interests focus on second language vocabulary acquisition and incidental and intentional language learning.

Kata Csizér is Associate Professor in the Department of English Applied Linguistics, Eötvös University, Budapest, and holds a PhD in Language Pedagogy. Her main field of research interest focuses on the sociopsychological aspects of L2 learning and teaching as well as second and foreign language motivation.

Peter I. De Costa is Assistant Professor in the Department of Linguistics and Languages at Michigan State University. His primary areas of research are identity and ideology. His work has appeared in *Language Learning*, *Language Teaching*, *System*, and *TESOL Quarterly*.

Robert DeKeyser is Professor of Second Language Acquisition at the University of Maryland and has a PhD from Stanford University. His interests include skill acquisition theory, the roles of implicit and explicit learning, individual differences, aptitude-treatment interaction, and study abroad. He is associate editor of *Bilingualism: Language and Cognition* and former editor of *Language Learning*.

Tracey M. Derwing is Professor Emeritus in Educational Psychology (TESL), University of Alberta, and Adjunct Professor in Linguistics at Simon Fraser University. Her interests include the interrelationships of comprehensibility and intelligibility with aspects of pronunciation, native speaker adjustments for L2 speakers, and the enhancement of immigrant settlement.

Jean-Marc Dewaele is Professor of Applied Linguistics and Multilingualism at Birkbeck, University of London. He is president of the *International Association of Multilingualism* and former president of the *European Second Language Association*. He is general editor of the *International Journal of Bilingual Education and Bilingualism*. He won the Equality and Diversity Research Award from the *British Association for Counselling and Psychotherapy* (2013) and the Robert Gardner Award for Excellence in Second Language and Bilingualism Research (2016) from the *International Association of Language and Social Psychology*.

Patricia A. Duff is Professor of Applied Linguistics at the University of British Columbia. Her research interests include language socialization and learning in multilingual contexts, the teaching/learning of English as a global language, and qualitative research methods in applied linguistics. She has published and lectured on these topics widely.

Rod Ellis is Research Professor at Curtin University, Australia, and Cheung Kong Scholar Chair Professor at Shanghai International Studies University. He was recently elected a fellow of the Royal Society of New Zealand. His published works include numerous articles and books on second language acquisition, language teaching and teacher education. His latest book is an edited collection, *Becoming and Being an Applied Linguist* (2015).

Susan Gass is University Distinguished Professor of Second Language Studies (SLS) at Michigan State University where she serves as director of the English Language Center and of the SLS Program. She has published widely in the field of second language acquisition including books on second language acquisition and research methods.

Beatriz González-Fernández is a PhD student of Applied Linguistics at the University of Nottingham. Her research interests include the processing, acquisition, and instruction of vocabulary in second languages. More specifically, her work examines how the different vocabulary knowledge components are acquired, in order to inform more systematic vocabulary pedagogy.

Laura Gurzynski-Weiss is Associate Professor in the Department of Spanish and Portuguese at Indiana University. She investigates interaction- and task-based classroom ISLA, feedback provision and perception, instructor cognition, research methodology, and the role of interlocutor and instructor individual characteristics as well as learner individual differences.

Sara Kennedy is Associate Professor of Applied Linguistics in the Department of Education at Concordia University in Montreal, Canada. Her research focuses on the learning, instruction, assessment, and use of second language speech, especially pronunciation.

Youjin Kim is Associate Professor in the Department of Applied Linguistics and ESL at Georgia State University. Her primary research interests include the role of interaction and

Contributors

individual differences in second language acquisition, task-based language teaching, and structural priming focusing both on adolescent and adult language learning.

Ute Knoch is the Director of the Language Testing Research Centre at the University of Melbourne. She has published in journals such as *Language Testing*, *Language Assessment Quarterly*, *Applied Linguistics*, *Assessing Writing*, and *Journal of Second Language Writing*. In her research she focusses on a variety of topics in the area of language assessment.

James P. Lantolf is Greer Professor in Language Acquisition and Applied Linguistics, and Director of the Center for Language Acquisition and the Center for Advanced Language Proficiency Education and Research (CALPER) at Pennsylvania State University. His research focuses on classroom language teaching, learning, and assessment.

Jongbong Lee is a PhD student in the Second Language Studies program at Michigan State University. He has taught English and Korean in various locations in Korea and the US. His research interests include second language writing, corpus linguistics, interactional feedback, and language assessment. He holds an MA in TESL from Georgetown University as well as a BA and an MA in English Language Education from Korea University.

Ronald P. Leow is Professor of Applied Linguistics and Director of Spanish Language Instruction in the Spanish and Portuguese Department at Georgetown University. His areas of expertise include language curriculum development, teacher education, instructed language learning, and cognitive processes. In addition to his numerous publications in prestigious journals, he has co-edited four books and authored one book (2015).

Shaofeng Li is Senior Lecturer at the University of Auckland, where he teaches postgraduate and undergraduate courses in second language acquisition and pedagogy and supervises Master's and PhD students. His research interests include language aptitude, corrective feedback, task-based instruction, and research methods.

Shawn Loewen is Associate Professor in the Second Language Studies and MA TESOL programs at Michigan State University. His research interests include instructed second language acquisition and quantitative research methodology. In addition to publishing in leading SLA journals, he has co-authored two books: *Key Concepts in Second Language Acquisition* and *An A–Z of Applied Linguistics Research Methods*. His sole authored book, *Introduction to Instructed Second Language Acquisition*, was published in 2015.

Roy Lyster is Professor of Second Language Education in the Department of Integrated Studies Education at McGill University in Montreal. He is author of two books: *Learning and Teaching Languages through Content* published in 2007 and *Vers une approche intégrée en immersion* published in 2016.

Kim McDonough is Professor and Canada Research Chair in Applied Linguistics at Concordia University. Her research interests include interaction and usage-based approaches to second language acquisition, structural priming, task-based language teaching, and collaborative writing.

Alison Mackey is Professor of Linguistics at Georgetown University (US) and Lancaster University (UK). Her research interests are in second language acquisition, specifically the role of input, interaction, feedback, and cognitive factors in L2 learning together with second language research design and methodology.

Susy Macqueen is Lecturer in the School of Literature, Languages & Linguistics at the Australian National University. She has experience teaching in primary, secondary, and tertiary educational contexts and has been involved in many language assessment projects, including test development projects for younger learners and evaluations of language learning programs.

Marije Michel is Assistant Professor at Lancaster University (UK) and holds a PhD from University of Amsterdam. Her research focuses on task-based performance, processing and assessment in second language acquisition and in particular the role of task complexity. Recently, she started using eye-tracking methodology to investigate attentional processes during second language (online) writing behavior.

Silvina Montrul is Professor of Linguistics and Spanish at the University of Illinois at Urbana-Champaign, with expertise in linguistic and psycholinguistic approaches to second language acquisition and bilingualism, language attrition, heritage languages, and the acquisition of Spanish. Her latest book is *The Acquisition of Heritage Languages*, published in 2016.

Hossein Nassaji is Professor and Department Chair in the Linguistics Department at the University of Victoria, Canada. His recent books are *Interactional Feedback Dimension in Instructed Second Language Learning* (2015) and *Teaching Grammar in Second Language Classrooms* (with Sandra Fotos, 2010). He is the winner of the Twenty-First Annual Kenneth W. Mildener Prize of Modern Language Association of America and the recipient of the 2012 Faculty of Humanities Award for Research Excellence, University of Victoria.

Bich Nguyen works as a research officer in the School of Education at Curtin University in Western Australia. Her research interests include world Englishes, systemic functionalism, Aboriginal education, and language teaching methodology. She has also been a second language teacher for 13 years.

Rhonda Oliver is Professor in the School of Education at Curtin University, Australia. Her work has appeared in a number of international and national journals including *Language Learning*, *Modern Language Journal*, *Studies in Second Language Acquisition*, and *TESOL Quarterly*. Her current research focuses on second language/dialect acquisition, particularly with children.

Carmen Pérez-Vidal is Associate Professor at the Universitat Pompeu Fabra, specializing in language acquisition with focuses on study abroad and immersion, and English linguistics. She has recently edited the volume *Language Acquisition in Study Abroad and Formal Instruction Contexts* published in 2014.

Luke Plonsky is Assistant Professor in the Department of Linguistics at Georgetown University, where he teaches courses in research methods and SLA. Publications in these and

Contributors

other areas appear in *Applied Linguistics*, *Language Learning*, and *Modern Language Journal* among other venues. He is associate editor of *Studies in Second Language Acquisition* and managing editor of *Foreign Language Annals*.

Charlene Polio is Professor and Associate Chair in the Department of Linguistics & Germanic, Slavic, Asian & African Languages at Michigan State University. Her main area of research is second language writing including the various research methods and the interface between the fields of L2 writing and second language acquisition. She has published in *TESOL Quarterly*, *Modern Language Journal*, the *Journal of Second Language Writing*, and *Studies in Second Language Acquisition* and is the co-author, with Debra Friedman, of *Understanding, Evaluating, and Conducting Second Language Writing Research* (2017).

Hayo Reinders (www.innovationinteaching.org) is Professor of Education at Unitec in New Zealand and TESOL Professor and Director of the doctoral program at Anaheim University, USA. His interests are in technology, learner autonomy, and out-of-class learning. He edits a book series for Palgrave Macmillan and is editor of the journal *Innovation in Language Learning and Teaching*.

Masatoshi Sato is Associate Professor of Applied Linguistics at Universidad Andres Bello, Chile. His research interests include peer interaction, corrective feedback, teacher education, CALL, and language testing. His work was recently published in *IRAL*, *Language Awareness*, *Language Learning*, *Language Teaching*, *SSLA*, *System*, and *Modern Language Journal*, and he co-edited *Peer interaction and second language learning: Pedagogical potential and research agenda* (2016).

Norbert Schmitt is Professor of Applied Linguistics at the University of Nottingham. He is interested in all aspects of second language vocabulary acquisition, description, and use. He has published six books and more than 50 journal articles on the topic.

Scott Sterling is Assistant Professor of Linguistics and TESL at Indiana State University. His main area of interest is in meta-research, focusing on the responsible conduct of research (RCR) in linguistics research.

Glenn Stockwell is Professor at Waseda University, Japan and holds a PhD from University of Queensland. He is co-author of *CALL Dimensions: Issues and Options in Computer Assisted Language Learning* (2006) with Mike Levy and editor of *Computer-Assisted Language Learning: Diversity in Research and Practice* (2012), and he has published numerous book chapters and articles in the field of CALL. He is editor-in-chief of the *JALT CALL Journal*, associate editor of *Computer Assisted Language Learning and Language Learning & Technology*, and on the editorial boards of *ReCALL*, *System* and the *CALICO Journal*.

Neomy Storch is Associate Professor in Applied Linguistics and the ESL Program Convenor in the School of Languages & Linguistics at the University of Melbourne. Her research focuses on issues related to ESL pedagogy, and in particular second language writing and the nature of peer interaction.

Pavel Trofimovich is Professor of Applied Linguistics in the Department of Education at Concordia University in Montreal, Canada. His research focuses on cognitive aspects of second language processing, second language phonology, sociolinguistic aspects of second language acquisition, and the teaching of second language pronunciation.

Lorena Valmori is a PhD candidate in Second Language Studies at Michigan State University. Her teaching experience has informed her research interests in motivational dynamics, identity, and emotions in second language learning. She has recently published her work in *Language Teaching*.

Bill VanPatten is Professor of Spanish and Second Language Studies at Michigan State University and affiliate faculty in Cognitive Science. He is widely known for his work in second language acquisition and instructed second language acquisition with particular focus on language processing and input processing.

Celia C. Zamora is a PhD candidate in Spanish Applied Linguistics from Georgetown University. Her dissertation looks to compare how heritage speakers and L2 learners of Spanish process +/- explicit feedback in a computerized task with concurrent data elicitation measures (Think Alouds). She serves as the Spanish Language Representative for the Center for Applied Linguistics (CAL).

Xian Zhang is Assistant Professor of Linguistics at University of North Texas. His research focuses on second language acquisition, including especially Chinese, language processing, cognitive neuroscience, lexical acquisition, and experimental research design.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Instructed Second Language Acquisition (ISLA)

An Overview

Shawn Loewen and Masatoshi Sato

What Is ISLA?

The field of instructed second language acquisition (ISLA) continues to be a growing sub-field within the discipline of second language acquisition (SLA) (see Nassaji, 2016). There are many similar concerns between the two fields, but the continued growth of second language (L2) learning and teaching, as a pedagogical, economic, social, and political activity, ensures that researchers, teachers, and learners continue to grapple with the practicalities of how best to acquire, learn, and teach an additional language.

There have been several attempts to define and describe the boundaries of ISLA (e.g., Ellis, 2005; Housen & Pierrard, 2005), with perhaps the most recent one found in Loewen (2015) in which he describes ISLA as

a theoretically and empirically based field of academic inquiry that aims to understand how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of an additional language.

p. 2

This definition focuses on several key aspects that will be explored further in this introductory chapter.

An Academic Field

An important starting point is that ISLA is an academic endeavor, meaning that it is based on a rigorous and scientific process of accumulating knowledge about L2 learning. To that end, theories and hypotheses have been and are being proposed about general or specific aspects of the L2 learning process (see VanPatten & Williams, 2015 for a recent overview of SLA theories); furthermore, these theories and hypotheses are investigated using data that researchers gather and interpret. Because researchers rely on specific skills and methods to research L2 learning (e.g., Larson-Hall, 2010; Mackey & Gass, 2015; Paltridge & Phakiti,

Shawn Loewen and Masatoshi Sato

2015), ISLA includes examination of research methodology, not because it necessarily has a direct impact on L2 learning (although in some cases it might, such as action research), but because research methods are lenses that provide information from specific epistemological perspectives. Consequently, methodology impacts the credibility and trustworthiness of research findings that ultimately inform pedagogical practice.

Systematic Manipulation

Another defining component of ISLA is the systematic manipulation of the learning environment and learning processes, which separates ISLA from what has been called, among other things, uninstructed or naturalistic L2 acquisition; in this, learners are simply surrounded by the target language but make no or little conscious effort to learn the language. Such scenarios might involve immigrants who are exposed to another language as they live in a wider social context, but who are not actively involved in learning the L2. Alternatively, uninstructed L2 learning might occur when expatriates who live and work in non-L1 contexts gain some knowledge of the local language, even though they are not concerned with achieving L2 proficiency. In both cases, the L2 may be “picked up” to a greater (in the case of immigrants) or lesser (for expatriates) degree, but the point is that there is no systematic effort by individuals to learn the L2 and/or by teachers/institutions to help develop the L2; rather, any L2 development results simply from exposure to the target language.

Instructional Contexts

The prototypical context for ISLA is, of course, the language classroom, which may take many different shapes: from introductory lessons for children in elementary school that aim to give kids a taste of an L2, to required university foreign language courses, to private language schools whose sole purpose is to promote L2 learning. However, it is important to point out that the physical classroom is not the only context of interest for ISLA because there is considerable L2 learning that occurs outside of the four walls of a classroom (Leow, 2015). For instance, the virtual L2 classroom is an increasingly popular L2 learning context, with both hybrid and fully online options (see Benson & Reinders, 2011). In addition, there are other circumstances, such as learner self-study, in which there is systematic manipulation of the learning conditions. For example, although autonomous learners may rely solely on authentic materials, in which case the level of manipulation is very low, learners generally use some type of study aid, such as books or computer programs or apps, to help them in their learning process. These materials, then, have been developed (i.e., manipulated) by individuals who presumably believe that the materials will be effective for L2 learning.

Another context that is included in ISLA is study abroad, even though the amount of manipulation may be minimal if students are placed in content classes taught in the target language and left to their own devices; however, many study abroad programs provide considerable structure for L2 learning. In such cases, learners are exposed to both intentional and incidental learning conditions (see Pérez-Vidal, 2014). As study abroad students interact in the broader target language context, they may not differ substantially from uninstructed learners; however, the mere fact that they have chosen to engage in study abroad indicates that they have altered their circumstances in an effort to gain more knowledge of the L2. Thus, although the amount of manipulation may vary, and it may be done by teachers, learners, or others (such as textbook designers), there is always at least some effort to acquire the L2.

Finally, it is important to point out that learning contexts may also affect the effectiveness of instruction because language instruction is a culturally bound endeavor, and while the fields of SLA and ISLA were primarily developed in North American and Western European contexts, the considerable importance of L2 instruction in other parts of the world has necessitated different perspectives on the classroom. In other words, it is necessary to conduct research in different learning contexts that may challenge existing ISLA theories or provide alternative perspectives. As an example, the different perspectives between task-based language teaching with its emphasis on student-centered activities (see Shehadeh & Coombe, 2012) and, in contrast, more teacher-centered educational cultures require ISLA researchers to consider how larger social, political, or ideological variables may affect the classroom (see Block, 2014).

Target of Manipulation

Another important consideration of ISLA are the mechanisms of learning, which include the processing and internalization of input; the restructuring, consolidation and storage of L2 knowledge; and the production of L2 output. However, not all learning mechanisms are of equal interest to ISLA researchers because some mental processes are not open to manipulation. For example, Universal Grammar (UG) or innatist perspectives of L2 acquisition are not primarily focused on instruction because arguably there is little that can be done to alter the makeup of the cognitive system. White (2015) states: “Clearly one cannot instruct L2ers as to UG-constraints (nor does anyone attempt to do so)” (p. 48). Similarly, the implicit processes that are involved in extracting patterns from input, as proposed by frequency- or usage-based approaches to L2 learning, are not generally influenced by L2 instruction, as Ellis and Wulff (2015) claim: “exemplar-based learning . . . is in large parts implicit . . . taking place without learners being consciously aware of it” (p. 76). Nevertheless, both innatist and frequency-based perspectives do have an interest in how the input that learners receive—which can be manipulated—affects the L2 learning process. In general, therefore, ISLA research is concerned with L2 learning processes that are hypothesized to be or have been found to be amenable to intervention.

Goals of Instruction

Having described ISLA in somewhat technical terms, it is important to consider, in more lay terms, its primary concern, which is: *what is the best way to learn and/or teach an additional language?* Implicit in this question is the notion that instruction can make a difference in L2 learning; however, the views about the amount of influence instruction can have on L2 learning range from minimal to extensive. For example, early theoretical views by Krashen (e.g., 1982, and more recently 2003), exemplified in a strong version of communicative language teaching (CLT), argue that instruction has little impact on L2 acquisition; instead, learners need to be provided with rich, authentic input in the classroom. Such views about the ineffectiveness of instruction, however, are in the minority, and most ISLA researchers, almost by definition, believe that instruction of some sort can positively influence L2 learning.

However, it is all well and good to say that L2 instruction is effective, but we also need to ask ourselves, *Effective for what?* In other words, what is the goal of L2 instruction? The goals of individual L2 learners or teachers may vary, but overall, the goal of many in the ISLA endeavor is for learners to develop communicative competence in the L2, that is the ability to use the L2 for communicative purposes (e.g., Littlewood, 2014). Of course, some

learners have other goals, such as gaining reading ability in the L2, learning phrases to help them on an upcoming trip, passing a L2 course required for their degree, or obtaining a good result on a standardized test to advance their careers. In other words, full proficiency or communicative competence may not be the goal. Nevertheless, if the goal of L2 instruction is often L2 proficiency, then we need to consider what precisely proficiency consists of, how to measure it, and what can bring it about.

Although there are different theoretical viewpoints about what constitutes L2 learners' linguistic knowledge, there is general agreement that not all knowledge is the same. On the one hand, there is what has been called explicit knowledge, declarative knowledge, or knowledge "about" language, all of which consist of information that learners are consciously aware of (DeKeyser, 2015; Rebuschat, 2013). Furthermore, this type of knowledge can be verbalized by learners and it can be reflected upon, although it may take the form of either lay terminology, such as "You need an *-s* because it is *he*," or more technical, metalinguistic descriptions, such as "third person singular *-s*." Another characteristic of explicit or declarative knowledge is that it is easily taught, in the same way as mathematic equations or historical dates. Teachers can present explicit information, often in the form of grammatical rules, and learners can commit them to memory. Subsequently, teachers can test to determine whether learners have retained this knowledge, and, if students have studied hard and have sufficient time to draw on their knowledge, they may do well on such tests.

However, the difficulty with explicit or declarative knowledge is that it is not readily available for use in spontaneous, real-time communication. For that, learners need to possess a type of knowledge that has been referred to variously as implicit knowledge, proceduralized knowledge, or knowledge "of" language, which is held unconsciously by the learner. In other words, learners are not aware of this knowledge, and they cannot verbalize it; however, learners are able to access it rapidly to communicate in spontaneous, real-time contexts. (Note, however, that it is possible for learners to possess both types of knowledge of the same linguistic feature.) The quintessential example of implicit knowledge is the knowledge that speakers have of their L1, especially before they receive any educational instruction about the language (via language arts or literature classes). When L2 learners ask L1 speakers why a specific utterance is grammatically or collocationally non-target-like, L1 speakers will often reply, "I don't know. It just sounds wrong." L1 speakers certainly know whether an utterance is acceptable in their L1, but they may not have the explicit knowledge of the linguistic rules to state why it is not acceptable. In sum, implicit knowledge is the primary contributor to communicative competence; therefore, it is the type of knowledge that many L2 learners wish to obtain and the type of knowledge that ISLA is primarily concerned with.

Specific language domains to which implicit knowledge can be deployed vary. Following the research focus in the field of linguistics, grammar has traditionally been the domain of ISLA research, with other linguistic areas receiving less coverage. However, that situation has changed over the past 20 years, with the increased emphasis on vocabulary, as well as pronunciation and pragmatics. Furthermore, one of the efforts of ISLA has been to provide a more integrated view of language and to consider ways in which theoretical concerns may apply across linguistic domains. So, for example, does the theoretical concern with explicit and implicit L2 knowledge, which has been primarily concerned with grammar, also apply to vocabulary, pronunciation, and pragmatics? Or are other theoretical perspectives more applicable? Although ISLA has been concerned with linguistic knowledge, there has also been a concern, especially among teachers and learners, with the language skills, especially productive skills. Consequently, some ISLA researchers conceptualize the goal of instruction in skill domains such as listening, reading, writing, and speaking.

Type of Instruction

While explicit knowledge (e.g., being able to recite grammatical rules) is relatively easy to gain and can be taught explicitly, implicit knowledge (e.g., being able to communicate in the target language accurately and fluently) is less amenable to instruction and often takes considerable time to develop. But if the goal of learners (and teachers and researchers) is implicit knowledge, how can this goal be achieved in the classroom? Can explicit knowledge be taught and then converted into implicit knowledge? ISLA scholars disagree on this point, which is referred to as interface positions. There are three perspectives: (1) the noninterface position maintains that the two types of knowledge are distinct and it is not possible for explicit knowledge to become implicit; (2) the weak interface position argues that under the right circumstances explicit knowledge may become implicit, but such conversion is not easy; and (3) the strong interface position claims that explicit knowledge can become implicit.

The reason that it is important to consider the relationship between explicit and implicit knowledge, from an ISLA perspective, is that it is important to know which types of manipulations (or instruction) are going to have an effect on which types of L2 knowledge. Within the last several decades, the investigation into this topic has been framed in terms of meaning-focused instruction and form-focused instruction. Meaning-focused instruction has its roots in the CLT movement, as put forward by researchers such as Krashen, who argued that the best way to bring about L2 communicative competence is by having learners communicate in the target language and that explicit instruction of linguistic forms (e.g., teaching grammar) has a detrimental effect on the development of communicative competence.

However, over time it became clear that meaning-focused instruction alone would not bring about the level of accuracy in L2 learner production that might be desired. Consequently, focus on form was put forward as a way of having brief attention to linguistic items during larger meaning-focused interaction (Long, 1996) in order to develop both accuracy and fluency in L2 learners. Long contrasted focus on form with focus on forms, the latter of which is the term he used for traditional, explicit language instruction. Over time, the terms *focus on form* and *focus on forms*, as well as *form-focused instruction* have been used somewhat differently by different researchers. Our current way of understanding of these terms (e.g., Loewen, 2015) is that form-focused instruction is a superordinate category that is commensurate with meaning-focused instruction; however, whereas meaning-focused instruction focuses exclusively on communication without any, or very minimal, attending focus on linguistic items or structures, form-focused instruction includes attention to linguistic form to varying degrees. Focus on form and focus on forms, then, are subordinate categories within form-focused instruction that reflect the amount of attention to linguistic structures in the instruction. In focus on forms, the primary focus is on linguistic structures, and instruction often follows a structural syllabus with different grammatical features being introduced in consecutive fashion. In contrast, focus on form describes instruction that is primarily meaning-focused, but includes brief attention to linguistic items as the need arises during communication. Sometimes, focus on forms and focus on form are used dichotomously to indicate two different types of instruction; however, it is perhaps more helpful to think of the two types of instruction as poles on a continuum, in which the ratio of attention to language form and meaning change proportionally.

So why does it matter how implicitly or explicitly language structures are addressed in instruction? Well, it goes back to the notion of what type of L2 knowledge teachers and researchers want learners to develop. There is a tendency for explicit instruction to result in

Shawn Loewen and Masatoshi Sato

explicit L2 knowledge, which tends not to be helpful in developing learners' communicative competence. Thus, the argument is that more implicit types of instruction, which have more emphasis on meaning and communication, are more suited for the development of implicit L2 knowledge. However, it is also the case that if instruction is too implicit, there may be no improvement in the accurate use of the targeted linguistic feature (as can be seen in fossilization of immersion learners). Currently, much ISLA research is ultimately concerned, either directly or indirectly, with the optimal combination of attention to language forms and language meaning in the classroom.

Having made a broad claim about the focus of ISLA research, it is important to acknowledge that there are numerous variables, both internal and external to the learner, which moderate and influence the effectiveness of instruction. Such individual differences are both interesting and challenging to ISLA researcher (as well as teachers and learners) who are trying to account for the effects of instruction. Learner-internal factors that have received considerable ISLA investigation include motivation, language aptitude, and foreign language anxiety (see Dörnyei & Ryan, 2015), while learner-external factors include the micro- and macro-social contexts in which learners find themselves (see The Douglas Fir Group, 2016). Furthermore, teachers' characteristics may affect the ultimate effect of instruction (see Borg & Sanchez, 2015).

In sum, this overview has attempted to provide an overarching framework for ISLA, while introducing the rich array of concerns and interests that comprise ISLA research. Given the diversity and complexity within the field, we refer the reader to the individual chapters included in the current handbook for specific theoretical foci, empirical references, and practical pedagogical suggestions.

About This Handbook

This handbook is the first collection of state-of-the-art papers pertaining to ISLA, with the purpose both to provide an overview of past ISLA research as well as to identify new and growing areas of interest. The handbook consists of 32 chapters (including the current chapter) written by 45 world-renowned experts and prominently emerging researchers in the field. Unlike many handbooks and encyclopedias, the entries are full-length articles detailing pertinent issues surrounding the respective topics. In addition, authors were asked to discuss updated research (as recent as 2017 publications) so that readers, both researchers and teachers alike, could be informed of current issues and cutting-edge pedagogical developments. We hoped to be comprehensive and inclusive in terms of topics but, at the same time, we are aware that such an endeavor never sees perfection.

The authors come from varying theoretical backgrounds precisely due to ISLA's cross-disciplinary nature (e.g., linguistics, psycholinguistics, psychology, sociolinguistics, technology, and education). Moreover, in order to reveal the complexity of L2 acquisition in instructional settings and to provide useful information to practitioners, we believed it was necessary to accumulate knowledge from differing perspectives. In this respect, we requested that authors share their expert opinions on their topics rather than merely surveying and summarizing existing research findings, with the result that each contribution constitutes a unique position paper. Also, we asked the authors to give a special attention to the *I* in ISLA by emphasizing pedagogical aspects and implications. As a result, we believe that each chapter serves three purposes: (1) providing updated literature and discussions of current issues; (2) sharing the authors' understanding of and approaches to the issues; and (3) providing direct links between research and practice.

Components of the Handbook

Each chapter starts with a Background section where the authors layout the framework for the topics. The following Current Issues section introduces theoretical and methodological issues that have been debated in the past, as well as those that are still being debated. Then, the authors elaborate the identified issues with empirical findings in the Empirical Evidence section. Importantly, the empirical evidence is discussed in order to support both the theoretical and pedagogical discussions. In the following Pedagogical Implications section (which occurs in all chapters except for those in Section I, focusing on theoretical issues, and Section VI, covering methodological concerns), the authors apply the empirical findings to instructional contexts. Finally, the authors conclude their chapters with the Future Directions section where they propose new research topics based on current studies and noticeable gaps in the research.

In addition to structuring each chapter in the aforementioned way, we asked the authors to include two types of call-out boxes. In Key Concepts boxes, the authors introduce and/or define concepts that are important to their topics. We hoped that the boxes would serve as a quick reference for a reader who may not be familiar with a particular topic. In the Teaching Tips call-out boxes, the authors offer practical pedagogical advice based on their research experiences. These call-out boxes can provide readers with a quick summary of some of the most important theoretical and pedagogical points in each chapter.

Topics in the Handbook

To achieve the goal of surveying research in the multifaceted discipline of ISLA, we divided the handbook into six sections.

- Section I: Second Language Processes and Products
- Section II: Approaches to Second Language Instruction
- Section III: Language and Instructed Second Language Acquisition
- Section IV: Instructed Second Language Acquisition Learning Environments
- Section V: Individual Differences and Instructed Second Language Acquisition
- Section VI: Instructed Second Language Acquisition Research Methods

It should be noted that in reality there is sometimes considerable and inevitable overlap between sections, and within chapters in a section. For example, Section I on L2 processes and products is more theoretical, but several of the chapters provide direct support for specific types of approaches to instruction in Section II. Additionally, different types of instruction (Section II) may be more or less relevant to specific aspects of language (Section III). Research both of learning and teaching environments (Section IV) and individual differences (Section V) require theoretical bases (Section I) and relate to instruction (Section II). Not to mention, research methodology (Section VI) is relevant to all research discussed throughout the handbook. The interconnection is a testimony of, again, the complexity of ISLA. Next we explain the main themes of each section and chapter.

Section I: Second Language Processes and Products

This section is probably the most theoretical and least directly applicable to the classroom; however, it is essential to understand the goals of ISLA—what is the result of ISLA—and how to achieve those goals. In Chapter 2, Robert DeKeyser dissects the issues related to

Shawn Loewen and Masatoshi Sato

L2 knowledge and skills (e.g., declarative/procedural, implicit/explicit, and automatized/controlled) and argues that the goal of ISLA is automatized procedural knowledge. He discusses different variables found to affect the development of such knowledge including the role of distributed practice, specificity of practice, and corrective feedback, all of which are relevant to classroom practice. Ronald P. Leow and Celia C. Zamora (Chapter 3) focus on mechanisms of L2 processing and type of L2 learning especially in relation to incidental/intentional learning. They caution that the construct of learning should be treated carefully in order to understand L2 processes (incidental/implicit vs. intentional/explicit) in instructional settings. In Chapter 4, Marije Michel discusses the result of L2 learning—complexity, accuracy, and fluency (CAF) in L2 production. The author provides a survey of CAF research and connects the findings to classroom assessment; she also calls for research to investigate the role of L2 production in the acquisition process. Finally, in Chapter 5, Neomy Storch adds a social perspective to ISLA. Based on sociocultural theory, the author argues for the inclusion of such perspectives in order to further our understanding of L2 learning processes and to better help teachers make pedagogical decisions (e.g., corrective feedback and group work).

Section II: Approaches to Second Language Instruction

This section explores different types of instruction that have been theoretically and empirically supported. In Chapter 6, Roy Lyster overviews a wide range of program types of content-based language teaching (CBLT) around the world. He makes a case for teaching language and content at the same time, with an emphasis on counterbalanced approaches to best assist the development of language skills in the classroom. Chapter 7 is devoted to task-based language teaching (TBLT). Rod Ellis first distinguishes TBLT from task-supported language teaching. He then shares practical suggestions as to what kinds of tasks to implement, how to implement them, and how to integrate tasks into a language curriculum. In Chapter 8, YouJin Kim summarizes research based on the interactionist perspective as a framework for ISLA. She offers suggestions as to how to enhance the effects of interaction, both between the teacher and learners and among learners, on L2 learning through corrective feedback, collaborative tasks, and learner training. James P. Lantolf and Xian Zhang (Chapter 9) discuss in detail a rather new pedagogical approach called concept-based language teaching. By reviewing sociocultural theory not only in relation to L2 education but to education in general, the authors introduce a Schema for the Orienting Basis of Action (SCOBA) for teaching a L2. In Chapter 10, Bill VanPatten provides a theoretical discussion of input processing and argues for processing instruction as a pedagogical intervention. He then suggests processing-oriented pedagogical interventions (POPIs) as a way of creating a mental representation of language based on input. Chapter 11 concerns a distinct yet important aspect of ISLA, that is, assessment. Ute Knoch and Susy Macqueen explain the concept of classroom-based assessment (CBA) and provide information pertaining to the timing and focus of assessment, as well as advice for individuals involved in the assessment process.

Section III: Language and Instructed Second Language Acquisition

This section addresses the different aspects of language that are the target of L2 instruction. First, Hossein Nassaji (Chapter 12) tackles arguably the most-investigated target in ISLA, namely, grammar. In reviewing major types of instruction (e.g., explicit/implicit,

focus-on-form/focus-on-form, input-based/output-based), the author reveals how they differentially assist different types of L2 knowledge. In Chapter 13, Kathleen Bardovi-Harlig focuses on pragmatics—the how-to-say-what-to-whom-when aspects of language. The author succinctly summarizes the challenges in teaching pragmatics or including it in a L2 program and provides empirical evidence that should be applied to L2 instruction. Chapter 14 concerns another linguistic target: fluency. Tracey M. Derwing discusses not only the processing aspects of fluency (or dysfluency) but also its social impacts. After reviewing pertinent research, the author introduces a variety of classroom activities designed to help learners develop fluency. Yet another important target in ISLA is pronunciation. In Chapter 15, in addition to discussing acoustic and perceptual aspects of L2 pronunciation, Sara Kennedy and Pavel Trofimovich emphasize the importance of considering pedagogical norms (e.g., nativeness versus intelligibility). The authors share their pedagogical perspectives by including various elements related to instruction of pronunciation (e.g., outside-class learning, teacher cognition, and computer-aided teaching). Chapter 16 concerns acquisition of vocabulary knowledge. Beatriz González-Fernández and Norbert Schmitt first summarize the historical background of vocabulary research in order to substantiate current pedagogical practices. Through the chapter, the authors provide the reader with useful pedagogical suggestions to increase both intentional and incidental exposure to target words in the classroom. Finally, Charlene Polio and Jongbong Lee (Chapter 17) take a different way of looking at L2 production, namely, L2 writing and its effects on the development of L2 knowledge. They too provide pedagogical suggestions based on updated research, especially related to written corrective feedback.

Section IV: Instructed Second Language Acquisition Learning Environments

This section acknowledges that ISLA is mediated by learning environments whereby target languages have different societal statuses and are learned differently due to different modes of communication. In Chapter 18, Yuko Goto Butler challenges some widely accepted ISLA norms (e.g., communicative competence, learner autonomy, and motivation) and argues that understanding L2 learning requires taking into account social/cultural perspectives, including the context in which the L2 is taught and learned. Focusing on Eastern Asian contexts, she proposes various contextually appropriate suggestions for L2 instruction. Another contextual variable that has been well investigated is study abroad. Carmen Pérez-Vidal (Chapter 19) discusses key differences between study abroad and study at home by focusing on contextual features (input and output opportunities), individuals' ability to make contact with the target language, and program features. She provides a useful list of program features that any language institute may want to consider for successful study abroad programs. In Chapter 20, Hayo Reinders and Glenn Stockwell overview the rapidly growing ISLA field of computer-assisted language learning (CALL). As technology develops and empirical findings from CALL research accumulate, the authors claim that CALL research can contribute to the development of SLA, as well as benefiting from it.

Section V: Individual Differences and Instructed Second Language Acquisition

This section addresses some of the individual differences that have been found to mediate SLA processes and the effects of instruction. In Chapter 21, Patricia A. Duff addresses social dimensions in ISLA (e.g., race, class, gender, sexuality, educational background,

immigration status, and ethnicity). She argues that people's perceptions and biases of social differences ultimately influence the outcome of SLA, and she proposes some ideas for teachers to consider in order to avoid negative impacts based on learners' social differences. Chapter 22, on the other hand, focuses on cognitive individual differences (i.e., language aptitude and working memory). Shaofeng Li reviews research examining the relationships between cognitive individual differences and types of instruction (e.g., explicit/implicit). He emphasizes that it is important, although challenging, to match learner types and instructional approaches in the classroom. Kata Csizér (Chapter 23) reports on self-related models and dynamics system theory in order to understand L2 motivation. Importantly, the author makes a direct and convincing connection between motivation research and classroom practice. In Chapter 24, Jean-Marc Dewaele provides a general review of psychological dimensions of ISLA including the higher order personality traits (the Big Five). In particular, he focuses on foreign language anxiety (FLA) and discusses how dynamically FLA is related to a web of personality traits and states. Laura Gurzynski-Weiss (Chapter 25) provides a perspective and research findings related to a necessary yet underinvestigated component of ISLA, that is, the teacher. In conceptualizing instructor individual characteristics (e.g., teachers' native language(s), years of teaching experience, educational background, engagement with research, etc.), the author establishes the significance of the research in relation to ISLA. Yet another individual difference that has been found to affect ISLA significantly is age. Rhonda Oliver, Bich Nguyen, and Masatoshi Sato (Chapter 26) collect a number of ISLA studies focusing on child L2 learners. While admitting methodological challenges in working with children, the authors lay out key similarities and differences between child SLA and adult SLA, including the need to be mindful of how the development of children's general cognitive abilities may influence L2 acquisition. The section ends with the topic of heritage language acquisition written by Silvina Montrul and Melissa Bowles (Chapter 27). As with the other individual differences, instructed heritage language learning presents unique variables and pedagogical challenges. Drawing on cognitive, sociocultural, and political perspectives, the authors discuss some important pedagogical questions, such as whether to include L2 learners and heritage language learners in the same classroom.

Section VI: Instructed Second Language Acquisition Research Methods

Finally, no academic discipline can advance without sound research. Consequently, this section attempts to capture the wide and developing range of research methods that are used in ISLA research. First, Luke Plonsky (Chapter 28) explains how important it is to increase objectivity, systematicity, and ease of analysis in advancing quantitative research, and he walks the reader through key decision-making points in conducting quantitative research. He also summarizes recent meta-analyses in ISLA. In contrast, qualitative methodology is explored by Peter I. De Costa, Lorena Valmori, and Ina Choi in Chapter 29. The authors proclaim that researching the mechanisms and conditions of L2 learning is insufficient to understand ISLA, and they propose that social dynamics (e.g., any semiotic resources available to learners in the classroom) need to be investigated. A series of exemplar studies helps the reader understand the nature and strengths of qualitative research methods. The following two chapters address the tension that exists concerning the validity of ISLA research. In Chapter 30, Alison Mackey reports on ISLA research conducted in the classroom setting. She succinctly summarizes data collection and analysis tools used in previous quasi-experimental studies and raises methodological challenges for classroom-based research. Kim McDonough

(Chapter 31), on the other hand, discusses research methodology and findings of common laboratory-based research, namely, structural priming, joint attention, and elicited imitation. The author calls for methodological rigor and validity in such experimental research methods, while acknowledging that a primary goal of such research is to inform classroom practice. The final chapter deals with research ethics, which is relevant and important for any type of research (Chapter 32). Susan Gass and Scott Sterling contend that following institutional guidelines (institutional review boards, or IRB) does not necessarily make a researcher ethical. On the contrary, researchers need to consider the possible consequences of their actions while conducting classroom studies. Particularly useful is the list of ethically focused scenarios that the reader can ponder. As the field of ISLA advances exponentially, ethical considerations are necessary to advance our research agenda.

Intended Audience of the Handbook

This handbook is intended for researchers, graduate students, upper-level undergraduate students, teachers, and teacher-educators who are interested in L2 learning and teaching. For undergraduate and nonthesis graduate students, the handbook provides an overview of the current state of the field of ISLA. Each chapter provides updated literature, which gives the reader an understanding of recent developments. For thesis graduate students or researchers, the chapters serve as useful reference points due to the thorough coverage of pertinent studies. Also, as the experts share their personal positions on various topics, readers may be able to situate themselves in the cutting-edge theoretical discussion. In the same vein, the research methodology section (Section VI) and the Future Directions segments in each chapter are useful for readers who are looking for a new research project.

For teachers and teacher-educators, theoretical debates or even research findings are sometimes inconsequential. Rather, what is often helpful for them is a list of potential pedagogical practices that they can employ in their classrooms. The pedagogical implications sections in each chapter provide such information. Also, the Teaching Tips boxes offer the reader quick suggestions while skimming through the chapter. We would like to stress that, unlike language textbooks and other pedagogically oriented volumes, the suggestions are based on empirical evidence on which teachers can confidently base their pedagogical decisions. We believe that, with nearly 40 years of investigation, ISLA research can and should contribute substantially to the classroom, and we hope that teachers find the pedagogical perspectives in this handbook relevant and useful.

Our Acknowledgments

First and foremost, we owe a debt of gratitude to the contributing authors, who were extremely long-suffering as they incorporated our sometimes demanding, persistent requests during the revision processes (*more, more, and more!*). All the chapters went through at least three rounds of revisions, and we humbly and sincerely hope that the authors share our view that our feedback was constructive and useful. Second, we would like to thank the anonymous reviewers of the initial book proposal. Our original thinking (e.g., the structure of the handbook) was pushed forward thanks to their useful suggestions. Also, we would like to thank our research assistant, Xuehong (Stella) He, who efficiently worked on formatting the manuscripts. Our sincere appreciation also goes to the Routledge editors, Leah Babb-Rosenfeld and Rebecca Novack, who enthusiastically supported the project from start to finish, and Kathrene Binag helped us through the production stage with amazing efficiency.

Shawn Loewen and Masatoshi Sato

Last but not least, we thank our respective families for their patient support especially when we were less than pleasant and reasonable. Without their support and input, this project would not have been possible.

References

- Benson, P., & Reinders, H. (Eds.). (2011). *Beyond the language classroom*. New York: Palgrave Macmillan.
- Block, D. (2014). *Social class in applied linguistics*. New York: Routledge.
- Borg, S., & Sanchez, H. (Eds.). (2015). *International perspectives on teacher research*. New York: Palgrave Macmillan.
- DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94–112). New York: Routledge.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York: Routledge.
- The Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, 100-S1, 19–47.
- Ellis, N. C., & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 75–93). New York: Routledge.
- Ellis, R. (2005). *Instructed second language acquisition: A literature review*. Wellington: Ministry of Education, New Zealand.
- Housen, A., & Pierrard, M. (2005). *Investigations in instructed second language acquisition*. Berlin: Mouton de Gruyter.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Portsmouth, NH: Heinemann.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.
- Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. New York: Routledge.
- Littlewood, W. (2014). Communication-oriented language teaching: Where are we now? Where do we go from here? *Language Teaching*, 47(3), 349–362.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic Press.
- Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design* (2nd ed.). New York: Routledge.
- Nassaji, H. (2016). Research timeline: Form-focused instruction and second language acquisition. *Language Teaching*, 49, 35–62.
- Paltridge, B., & Phakiti, A. (Eds.). (2015). *Research methods in applied linguistics: A practical resource*. London: Bloomsbury.
- Pérez-Vidal, C. (Ed.). (2014). *Language acquisition in study abroad and formal instruction contexts*. Amsterdam: John Benjamins.
- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63, 595–626.
- Shehadeh, A., & Coombe, C. (Eds.). (2012). *Task-based language teaching in foreign language contexts: Research and implementation*. Amsterdam: John Benjamins.
- VanPatten, B., & Williams, J. (Eds.). (2015). *Theories in second language acquisition: An introduction* (2nd ed.). New York: Routledge.
- White, L. (2015). Linguistic theory, universal grammar, and second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 34–53). New York: Routledge.

Section I

Second Language Processes and Products



Taylor & Francis

Taylor & Francis Group
<http://taylorandfrancis.com>

Knowledge and Skill in ISLA

Robert DeKeyser

Background

When somebody asks “How many languages do you know?” what does the word ‘know’ mean? Does it mean the same for the one asking the question, who has never learned a second/foreign language, as for the interlocutor, who has learned several? When somebody else asks the same person “How many languages do you speak?” will that elicit the same answer? Will ‘speaking a language’ mean the same for both interlocutors in this case? Everybody who reads these sentences has probably learned at least one additional language, has probably been asked questions like these, and therefore realizes that the questioner and the interlocutor probably understand the word ‘know’ or ‘speak’ in very different ways. ‘Knowing’ or ‘speaking’ a language is a complex concept. The more experience we have learning languages, and the more research we carry out on language learning, the more we realize how complex.

While any beginning foreign language learner realizes that knowledge of vocabulary, of grammar, and of pronunciation are very different things, most of the distinctions that we constantly make in second language acquisition (SLA) research are less obvious, however, and are a frequent source of confusion and frustration, even for the researchers themselves. Dichotomies abound: implicit/explicit knowledge, declarative/procedural knowledge, incidental/intentional learning, instructed/naturalistic learning, inductive/deductive learning, and item/rule learning, to name just the most common ones. What exactly do these distinctions mean and how do they relate to each other? The question gets even trickier when we start asking about the relationship between implicit/explicit knowledge, implicit/explicit learning, and implicit/explicit teaching. These questions, however, are the most important ones of all from an applied perspective: if we want the learner to end up with a certain type of knowledge, what does that imply for how learning should happen, and in turn for how instruction should proceed?

In the past, languages were most often taught like any other subject matter: grammar rules were explained, vocabulary lists were memorized, and then the student was tested on this knowledge through fill-in-the-blanks exercises or at best a brief translation. Skill in using the language for interpersonal communication was rarely an issue and, not surprisingly, was not acquired. The distinction between knowledge and skill, then, is one that most foreign language learners are painfully familiar with, to the extent that students will sometimes

Robert DeKeyser

rebel against the perceived emphasis on knowledge and the lack of ensuing skill by saying “We don’t want to learn grammar, we want to learn how to speak.” The student saying this is mixing up several distinctions from the point of view of the researcher, but the point is well taken: given that nowadays the goal of most language teaching IS skill, what does that imply for the kind of knowledge that should be attained eventually, and for the best path to get there? How does this perhaps most basic distinction of all, between knowledge and skill, relate to the other dichotomies mentioned?

First of all, skill *is* a form of knowledge. When in everyday language we say somebody knows a lot, we think of facts and figures, not of that somebody being good at basketball, at singing, or at chess. Yet, in more technical terms, the accomplished basketball player possesses a form of knowledge permanently stored in memory and drawn on constantly for executing those skills: procedural knowledge, as opposed to the declarative knowledge that we usually designate as knowledge. Sometimes declarative knowledge is called ‘knowledge that’ and procedural knowledge is ‘knowledge how.’ This distinction is easy to misunderstand, however. When a learner knows that an English verb takes a final *-s* in the third person singular, one could say this learner knows how to conjugate a verb, or when to use that final *-s*, but this knowledge is not procedural unless the learner has executed the mental act of selecting that morpheme under the right conditions many times, and has therefore learned a behavior instead of knowledge about a desirable behavior. A slightly more elegant definition of declarative versus procedural knowledge, then, is the following: “Declarative representations are objects of thought, whereas procedural representations provide the (cognitive) actions to work upon these objects” (Gade, Druey, Souza, & Oberauer, 2014, p. 174).

Procedural is not exactly the opposite of declarative. The famous case of patient H.M., whose memory was selectively impaired after a brain trauma, led to a large number of studies on how different aspects of memory are dissociated, and the main distinction is between declarative and nondeclarative. At least four kinds of nondeclarative memory are distinguished: procedural, priming, simple classical conditioning, and habituation (Corkin, 2013; Henke, 2010; Squire & Wixted, 2011). For the purpose of explaining skill acquisition, however, the main kind of nondeclarative knowledge is procedural knowledge, as opposed to declarative knowledge.

This brings us to the relationship between the declarative/procedural and the explicit/implicit distinction. For many practical purposes, the two dichotomies are equivalent, but from the perspective of cognitive neuroscience, they are not. Explicit knowledge is knowledge one is aware of, and implicit knowledge is knowledge without awareness (see A. Reber, 2003; P. Reber, 2013; Rebuschat, 2013; Williams, 2009). Declarative knowledge is mostly explicit for our purposes (classroom language learning), but can be implicit (as in the Chomskyan concept of grammatical competence). Explicit knowledge, however, is considered by many to be necessarily declarative (Paradis, 2009; Ullman, 2015).

In most forms of second language instruction, even today, the learning of grammar starts as both declarative and explicit. As a result, learners know what they should do and are aware of what they know, but are not able to do what they know they should do unless they are focused on form and have enough time to draw on their declarative knowledge and act upon it using high-level, that is very abstract, all-purpose procedures. As a result of practice they become better at putting their knowledge to use, using it more correctly, more easily, more frequently, in a wider variety of contexts. Sometimes this process is called automatization in a broad sense, but more technically what happens is first developing procedural knowledge (which happens relatively fast), and then automatizing it (which takes a very long time and for most learners and most structures probably never reaches asymptote).

Highly automatized knowledge is usually characterized as unintentional, uncontrollable, unconscious, efficient, and fast, but not all psychologists see all these characteristics as essential to the concept of automaticity (for a very thorough discussion of automaticity and how its various aspects relate to each other, see esp. Moors & De Houwer, 2006).

The role of practice in getting to a sufficiently high level of automatization to enable second language use that is both fluent and almost completely accurate is one of the most central topics in instructed second language acquisition (ISLA). Practice itself comes in many varieties. In the broadest sense it simply means using the language for communication; this is the meaning of the term in contexts like “you have to go abroad to get enough practice.” In the narrowest sense it means deliberate practice, which Brown, Roediger, and McDaniel (2014, p. 183) characterize as follows: “If doing something repeatedly might be considered practice, deliberate practice is a different animal: it’s goal-directed, often solitary, and consists of striving to reach beyond your current level of performance.” In other words, acquisition of skill is not a by-product of the practice here; it is the one and only goal. All drills fall into this category of practice, whether they be mechanical, meaningful, or communicative (Paulston, 1972). Between these two poles of activities, those meant exclusively for learning and those that can have (incidental) learning as a by-product, there are many intermediate varieties: role plays, skits, scenarios, fill-in-the-gap tasks, picture descriptions, and a wide variety of other tasks.

While nobody would contest the need for practice of some kind, one of the most controversial issues in language teaching concerns what kind of practice is best. The answer depends, of course, on a variety of factors such as the nature of the knowledge before practice, the kind of skill desired as a result of practice, and the time and resources available. What constitutes good practice activities will be one of the core issues discussed in the next session.

Key Concepts

Declarative knowledge: Knowledge of facts (semantic memory) and events (episodic memory); usually consciously accessible and often verbalizable, but not necessarily; sometimes called knowledge THAT as opposed to knowledge HOW.

Procedural knowledge: Knowledge that can only be performed, such as how to swim, do mental arithmetic, or speak fluently; sometimes called knowledge HOW as opposed to knowledge THAT.

Proceduralization: The process of creating procedural knowledge by incorporating elements of declarative knowledge into broader preexisting procedural rules. This takes place when learners repeatedly engage in a task that calls on the same declarative knowledge.

Automatization: The gradual improvement in speed, error rate, and effort required that characterizes performance on repeatedly practice tasks; this improvement is made possible by restructuring the components of procedural knowledge and not by merely speeding up its use.

Deliberate practice: Activity of repeatedly engaging in a behavior in order to become better at it (as opposed to the incidental practice that comes with activities engaged in repeatedly for work or personal routines).

Skill specificity: The specialized nature of procedural knowledge, which causes it not to be directly transferable to other skills, in particular from comprehension to production and vice versa, but only indirectly via declarative knowledge.

Transfer-appropriate processing: Processing that has enough elements in common with the context of transfer for this context to activate the memory traces from this processing.

Current Issues

As questions about the kinds of knowledge that are desirable and the way to acquire them are at the core of our field, it is not surprising that they are at the intersection of a number of intense debates. How much implicit learning is possible and how much explicit learning is necessary? What kind of knowledge results from these kinds of learning? How can that knowledge change over time? What kinds of experiences can lead to those changes? This section will serve to flesh out these questions a bit more and to provide some tentative and interrelated answers. The next section will then provide a bird's-eye view of the wide array of empirical research on these various questions and will be able to treat each topic more independently, once the broad picture has been presented here.

As we saw in the previous section, being skilled at something means one has the requisite procedural knowledge. Mere knowledge of grammar rules and vocabulary does not suffice; one needs to be able to use knowledge fast and accurately, and that means one needs a large store of procedural knowledge, including that required to fill in its own gaps by drawing on the necessary bits of declarative knowledge and incorporating them seamlessly into one's communicative behavior. The ability to do the latter rests on higher-level procedures, often referred to as communication strategies or strategic competence (e.g., McNamara, 1995). Much strategic behavior stands out because it makes the L2 speaker fall back on circumlocutions, avoidance strategies, L1 transfer, or gestures, which do not necessarily improve accuracy and have been the object of a body of research in their own right (e.g., Kasper & Kellerman, 1997; Lafford, 2004; Macaro, 2006), but skilled L2 speakers will often be able to fill the gaps in their procedural knowledge by drawing very efficiently on declarative knowledge without uttering anything that can be detected as nonnative or even nonfluent. High levels of fluency leave enough mental resources to plan ahead, detect possible sources of nonfluency or nonaccuracy, and avoid them by searching efficiently for alternative procedures, including procedures that call on small chunks of declarative knowledge.

Such high levels of fluency require not just procedural knowledge, but automatized procedural knowledge. This does not imply the use of these rules should be entirely automatic; automaticity is a graded concept (DeKeyser, 2001; DeKeyser & Criado-Sánchez, 2012; Segalowitz, 2010; but see Paradis, 2009 for a dissenting opinion), perhaps even multicomponential (Moors & De Houwer, 2006), and for most skills and subskills we never reach the asymptote in the learning curve for error rate and reaction time that would mark the end point of automatization—not even for mental arithmetic, typing, driving a car, or speaking our native language.

Few researchers would disagree with the gist of the previous two paragraphs even though they may use different terminology and put different emphases. Even assuming one agrees with all of this, however, many questions remain. For instance, does fully automatized knowledge mean implicit knowledge? Hardly any research has addressed that question, but Suzuki and DeKeyser's (submitted) findings suggest that automatized knowledge can become implicit. On the other hand, even fairly highly automatized procedural knowledge certainly is not necessarily implicit (Suzuki & DeKeyser, 2015). Needless to say, if automatized implicit knowledge is not necessarily implicit, procedural knowledge in more initial stages of development is even far less likely to be so. Some claim all procedural knowledge is implicit (e.g., Ullman, 2015), but given how procedural knowledge develops by engaging in the target behavior while drawing on explicit declarative knowledge, and how even highly automatized knowledge does not always seem to be implicit, it may not be desirable to draw such a strict line.

The claim that procedural knowledge is necessarily implicit would follow logically from the claim made by some that declarative knowledge cannot be proceduralized or automatized (Paradis, 2009; Ullman, 2015). This point of view is known in applied linguistics as the ‘noninterface position’ (even though most proponents of that position use the term ‘explicit’ instead of ‘declarative’). The noninterface point of view is perhaps most strongly associated with Krashen (1982, 1985), and endorsed by a number of researchers (e.g., Truscott, 1998). The noninterface position, however, seems based on an overly radical interpretation of ‘interface,’ and tied to the somewhat misleading wording about declarative knowledge getting ‘converted’ or ‘transformed’ into procedural knowledge. This terminology may be seen as implying that the more procedural knowledge there is on a given point, the less declarative knowledge there is, which often is not true (DeKeyser, 2009). Nor is it the case that declarative knowledge somehow moves from the parts of the brain where declarative knowledge seems to be stored, that is, the hippocampus and the temporal cortex, to the areas where procedural knowledge is stored, that is, the basal ganglia and the frontal cortex (Henke, 2010; Ullman, 2015). What does seem to be the case, however—and this is crucial—is that declarative knowledge allows learners to engage in the target behavior (e.g., using a morphosyntactic rule in communication), and by drawing on this declarative knowledge repeatedly to engage in this behavior repeatedly, forming procedural knowledge, establishing a habit after some repetition, and then gradually automatizing this habit, and perhaps eventually (for some structures in some people) implicit knowledge. It’s not like one brain circuit ‘infects’ the other, but rather that one memory system enables behaviors that lead to the gradual establishment of another memory system (Hulstijn, 2002; see also DeKeyser, 2015; Paradis, 2009). Such development of memory in one brain area by drawing on related memory in another area is nothing unusual, and is also seen, for example, in the development of declarative knowledge over time or procedural knowledge over time (e.g., Chein & Schneider, 2005; Hill & Schneider, 2006; Kelly & Garavan, 2005; Opitz & Friederici, 2003). Speaking about the development of declarative knowledge in particular, Squire and Wixted (2011, p. 273) state “The idea is not that memory is literally transferred from the hippocampus to neocortex but that gradual changes in the neocortex increase the complexity, distribution, and connectivity among multiple cortical regions.”

From the point of view of ISLA research, the declarative-procedural-automatized distinction is more important than the implicit-explicit distinction. The fact that it is very hard to find a pure measure of implicit knowledge (Jiang, 2012; Rebuschat, 2013; Suzuki & DeKeyser, 2015) proves that implicit knowledge and highly automatized knowledge are functionally equivalent, in the sense that they cannot be distinguished in communicative interaction; it takes very carefully calibrated laboratory experiments to distinguish the two. All learners, therefore, would be perfectly happy with highly automatized procedural knowledge, and their teachers and employers too, leaving it to psycholinguists to worry about the degree of implicitness.

Proceduralization and automatization, then, are essential if the learner is to become perfectly fluent (in the broad sense of able to communicate at normal speed with a high degree of accuracy), which means systematic practice is crucial. The role of various kinds of practice in the development of proficiency is the subject of a sizeable chunk of the ISLA literature. Most of it focuses on accuracy, some on fluency (in the narrow sense of being able to speak smoothly, at a normal speed, without many hesitations and pauses), and far less deals with complexity (but see the special issue of *SSLA* 2016 on the topic). Moreover, very little ISLA research addresses proceduralization and automatization directly, even

Robert DeKeyser

though the combination of high degrees of accuracy, fluency, and complexity presupposes an advanced degree of automatization (for more on complexity, accuracy, and fluency, see Michel, this volume).

The research that does focus on proceduralization or automatization, however, has largely confirmed the picture from skill acquisition theory: (1) error rate and reaction time both tend to decline following a power function (a steep decline at first, which rather suddenly turns into a very slow decline; cf. DeKeyser, 1997; Rodgers, 2011); (2) once knowledge has been proceduralized, it is very skill-specific, which means in particular that production practice tends to be good for skill in production, but less for skill in comprehension, and the other way around (e.g., DeKeyser, 1997; Li & DeKeyser, in press; cf. Towell, 2012). Research conducted in the framework of processing instruction usually finds far less specificity, and often more transfer from comprehension practice to production outcomes than the other way around (Shintani, Li, & Ellis, 2013), which may be due to methodological features of that line of research (such as the noncommunicative practice and testing formats and the choice of linguistic features), or may indicate that the knowledge measured in that research is not truly proceduralized, or both. The fact that more communicative implementations of production exercises in a processing instruction framework have come to more positive conclusions about the role of production practice (cf. DeKeyser & Prieto Botana, 2015) certainly points in that direction.

Another big topic in the area of practice is not about its nature but its distribution over time. The effect of distributing practice has been studied in cognitive and educational psychology for many decades, and the general consensus is that distributing practice is usually beneficial (Carpenter, Cepeda, Rohrer, Kang, & Pashler, 2012; Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Rohrer, 2015). Several questions remain, however, about the ideal spread over time, the extent to which the benefits apply to different kinds of knowledge, and the ultimate reason for distribution effects (Toppino & Gerbier, 2014). There is considerable evidence in the psychology literature for an ideal ratio between distribution of practice and the time of testing (Cepeda et al., 2006), but that body of research is mostly about paired-associate learning (word lists and so on), and clear evidence for this ratio has not been found yet in SLA research. Some have even suggested that massed practice is better when one looks at the L2 curriculum as a whole instead of very specific elements of grammar and vocabulary. An intense summer course in L2 may lead to more proficiency than an equivalent amount of time spread over several semesters (Serrano, 2011).

The issue of practice is intimately linked with that of corrective feedback (CF); it is hard to separate one's views on skill acquisition and practice in general from one's perspective on feedback. Here too, a distinction has often been made between implicit and explicit. Lyster, Saito, and Sato (2013), for instance, put all forms of CF on a continuum from implicit (clarification requests, conversational recasts) to explicit (providing metalinguistic clues, explicit correction with metalinguistic explanation). It is important to realize that the way feedback is given does not necessarily correspond to the processes of learning or the nature of knowledge resulting from them. Regardless of how implicit a recast may be, if the learner notices its form and its corrective intent, then there is awareness of what is being learned, that is, there is explicit learning, and the knowledge immediately resulting from it is explicit. A considerable number of studies have compared implicit and explicit forms of error correction, and several meta-analyses have been conducted on this point. Lyster and Saito (2010) showed that prompts (providing negative feedback) were more effective than recasts (providing mostly positive feedback), with the effect of explicit correction (which provides both negative and positive feedback) not distinguishable from either prompts or

recasts. Mackey and Goo (2007), however, found a bigger effect for recasts, while Li (2010) stressed that implicit feedback did equally well or better in delayed testing. As Lyster and Saito included only classroom studies in their meta-analyses, while Mackey and Goo as well as Li also included laboratory studies, the difference in findings may well be attributed to this difference in contexts from which the studies in the meta-analyses were sampled, as Lyster and Saito suggest. What all three meta-analyses have shown, however, along with the one by Russell and Spada (2006), is that these various forms of CF are effective (if not always efficient), judging from both immediate and delayed testing.

What none of these meta-analyses address is to what extent recasts were successful for grammar specifically, and in particular for grammar not yet covered explicitly in class. One may expect that the success of recasts depends on how salient they are, and it has been shown that correction of morphosyntax, compared that in other linguistic domains, is both less likely to be noticed and less likely to be interpreted correctly when noticed. Mackey, Gass, and McDonough (2000), for instance, showed that corrections for lexis were most likely to be interpreted as such, while this interpretation was the least likely for morphosyntax, with pronunciation in between. As grammar is less salient than vocabulary to most learners, and as correction of grammar requires much more inferencing from the specific correction to the underlying rule that was violated, it is less likely that a grammar recast will lead beyond successful uptake. Intake in the sense of internalizing the underlying rule is still less likely, unless perhaps the rule is fairly well known to the learner already, in which case the recast, even more clearly than in general, serves to trigger explicit processes. Such recasts, as well as prompts and explicit correction, again assuming previous familiarity with the rules, thus serve to steer the proceduralization of declarative knowledge. As the skill acquisition literature has shown, engaging in the target behavior while the relevant declarative knowledge is kept in mind is essential to guide proceduralization (DeKeyser, 2015). This interpretation is very much in line with the findings in Sato and Lyster (2012), where CF by peers was found to be effective in improving accuracy without jeopardizing fluency, because the peer corrections made students monitor their declarative knowledge in production, and thus proceduralize and automatize it (see more detail on this study in the next section).

Empirical Evidence

The most central issue discussed in the previous section is the interface issue, especially in a broad sense: to what extent and how do declarative, procedural, automatized, and implicit knowledge interact with each other, feed into each other, or compete with each other? Thorough theoretical discussions on this point can be found in DeKeyser (2009), N. Ellis (2005), Hulstijn (2002), Lyster and Sato (2013), Morgan-Short (2012a, 2012b), Paradis (2009), Robinson, Mackey, Gass, and Schmidt (2012), and Spada and Lightbown (2012). When it comes to empirical research on the interface issue(s), or rather this set of related issues, it is very hard to draw the line between what is highly relevant, somewhat relevant, or not at all relevant to answering the interface question. The reasons for that are many. First, there is the conceptual overlap between dichotomies such as declarative/procedural, explicit/implicit, intentional/incidental, and controlled/automatic. For some questions in some contexts the concepts overlap enough that the distinctions do not matter, but for others they do. Second, even when the same dichotomies are used, and with the same definitions, their operationalizations in both treatments and outcome measures still vary widely. Third, two perennial problems in SLA research are perhaps felt even more strongly in this area: the overgeneralization of findings over structures, contexts, age ranges, and other individual differences;

Robert DeKeyser

and the unavoidable trade-off between internal and external validity. Finally, almost no research takes a longitudinal perspective that goes further than a few treatment sessions and a delayed posttest, even though most interface questions are inherently about development over time, often over a period of years.

For the implicit/explicit instruction distinction, we do have three meta-analyses of studies that have operationalized the distinction in a fairly consistent way. Both Goo, Grarena, Yilmaz, and Novella (2015) and Norris and Ortega (2000) found a clear advantage for explicit teaching over implicit teaching, even though they are based on a rather different set of studies: Goo et al. used 34 primary studies, only 11 of which overlapped with Norris and Ortega's dataset of 49. A third meta-analysis, by Spada and Tomita (2010), including 30 studies, 10 of which also figured in Norris and Ortega's meta-analysis, hypothesized an interaction between explicitness and complexity of structure, but again found a main effect in favor of explicit teaching. Two further meta-analyses found an advantage of explicit over implicit error correction: Li (2010) based on 33 studies and Lyster and Saito (2010), based on 15. It should be noted that in Li's analysis the advantage of explicit correction had faded in delayed testing, but on the other hand this last comparison was based on a small number of studies, and the effect sizes for explicit and implicit CF were not significantly different.

Given this overwhelming evidence in favor of explicit teaching and error correction from dozens of studies and five meta-analyses, and given the large overlap in practice between explicit/implicit and declarative/procedural (at least in the L2 teaching context), one may be tempted to assume that this literature also strongly suggests the advantage of initially acquiring explicit declarative knowledge. Two important caveats should be kept in mind, however. First, as Doughty (2003) pointed out in reaction to Norris and Ortega's meta-analysis, and as is still the case today, the outcome measures in almost all the primary studies concerned have been heavily biased toward explicit knowledge, which could in part explain the advantage found for explicit teaching. Second, the primary studies varied widely in context (laboratory, classroom, and group setting), stage of learning, and teaching activities. Therefore, this whole body of research says very little about skill acquisition: did the learners in these studies have declarative, (partially) proceduralized, or (partially) automatized knowledge? If they did, how did they acquire it? We don't know, certainly not for these studies in aggregate.

Studies do exist, however, that were specifically framed in skill acquisition terms. They deal with the nature of practice that leads to proceduralization, the specificity of procedural knowledge, the role of error correction, skill development during study abroad, and the need for transfer-appropriate practice.

DeKeyser (1997) and Ferman, Olshtain, Schechtman, and Karni (2009), through a longitudinal design, and Rodgers (2011), through a cross-sectional design, showed in detail how the use of morphosyntactic structures that had been taught explicitly showed a gradual decline in error rate and reaction time with practice, provided the declarative knowledge had been acquired. DeKeyser (1997) and Ferman et al. (2009) also showed, more specifically, a decrease in the form of a power function, which is characteristic of the acquisition of both psychomotor and cognitive skills (Newell & Rosenbloom, 1981; cf. DeKeyser, 2001). The fast decline in the first part of the learning curve is often interpreted as proceduralization, while the slow decrease in the second part is seen as evidence of automatization.

It should be pointed out, however, that a mere decline in error rate and reaction time is not necessarily evidence for automatization in the narrowest sense. Segalowitz and Segalowitz (1993; cf. also Segalowitz, 2010; Segalowitz, Segalowitz, & Wood, 1998) argued that automatization in the strict sense implies a restructuring of cognitive processes that should be reflected in a decrease in the coefficient of variation (the ratio of standard

deviation over mean, reflecting a change in processes, not just a speed-up). This view is still being debated, however (Hulstijn, van Gelderen, & Schoonen, 2009; cf. also DeKeyser & Criado-Sánchez, 2012), and empirical results have been mixed (see e.g., Lim & Godfroid, 2015; Rodgers, 2011).

On the other hand, some neuroscience work has documented the expected shift in the nature and location of activity in the brain as predicted by skill acquisition theory: Bowden et al. (2010, 2013) documented a LAN/P600 response to violations of inflectional morphology in Spanish L2 by native speakers and advanced L2 learners, but not intermediate learners. Morgan-Short, Sanz, Steinhauer, and Ullman (2010) showed a similar change in the processing of gender agreement. Morgan-Short, Faretta-Stutenberg, Bill-Schuetz, Carpenter, and Wong (2014), using individual differences measures of aptitudes for declarative and procedural memory, showed a shift from reliance on declarative to procedural memory capacity in participants learning the morphosyntax of an artificial language. Tanner, McLaughlin, Herschensohn, and Osterhout (2013), also using an individual differences approach, provided evidence for a shift in processing mechanisms for subject-verb agreement in German L2.

A more global approach to proceduralization was taken by de Jong and Perfetti (2011), who showed that repeated performance of an oral narration under increasing time pressure (4–3–2 minutes) led to increased fluency (as measured by articulation rate, phonation rate, length of fluent runs, length of pauses), which they interpret as evidence of proceduralization. They did not look at accuracy, however. Thai and Boers (2016), on the other hand, not only showed increased oral fluency as a result of the 4–3–2 task, but also that it was the repetition alone that mattered, not the increased time pressure, because a comparison group with a 2–2–2 version of the task showed the same improvement in fluency, and without showing the detrimental effect on accuracy that they found in the 4–3–2 condition.

An even more global approach is found in some of the literature on study abroad. On the one hand, studies such as O'Brien, Segalowitz, Freed, and Collentine (2007) and Segalowitz and Freed (2004) showed that fluency in the use of previously acquired declarative knowledge increases considerably during study abroad, depending on factors such as linguistic readiness (initial proficiency) and cognitive readiness (lexical access, attention control, and phonological short-term memory). Kahng (2014) also argued that her English as a Second Language (ESL) data from Korean L1 speakers at two levels of proficiency showed more reliance on declarative knowledge at the lower proficiency level. On the other hand, DeKeyser (2007, 2010a) showed that acquiring new declarative knowledge or trying to proceduralize shaky declarative knowledge was very difficult for students on a program abroad in Argentina, and that the most substantial gains in fluency were made by the students who monitored their declarative knowledge most consistently. Golonka (2006) also showed that the strongest predictors of proficiency gains during a stay in Russia were previous knowledge (grammar and vocabulary) and monitoring (self-correction and sentence repair). Together these findings certainly suggest that in spite of the change in context from the classroom to study abroad, it is declarative knowledge and practice to proceduralize and automatize it that determine how much fluency is gained, not a completely independent process of acquiring procedural (let alone implicit) knowledge “from scratch,” without drawing on declarative knowledge.

Even students returning from study abroad with substantially increased fluency, however, do not necessarily perform better than before in a variety of contexts, including the classroom. The conversational skills acquired in specific contexts abroad are not the ones need for debates, business negotiations, essay-writing, and so on, and even when it comes to

conversational skills, small differences in context can interfere with transfer of skill. When skills are quite different, for example, listening versus speaking, only the skill that is practiced (extensively) is proceduralized (and automatized). Transfer to the opposite skill happens via declarative knowledge (Singley & Anderson, 1989); the initial declarative knowledge that was available before one of the skills was practiced needs to be exploited again for creating the opposite skill. Therefore performance in the other skill is characterized by a much higher error rate and reaction time.

For the domain of grammar, this skill specificity of practice was demonstrated by de Jong (2005) and DeKeyser (1997). DeKeyser showed how learners of an artificial language performed better on specific morphosyntactic rules if they had practiced that rule in the same skill they were tested on: they did better on a comprehension test for rules practiced in comprehension and on a production test for rules practiced in production, for both reaction time and error rate. De Jong showed that learners who had practiced adjective-noun agreement in comprehension but not in production did even worse on a production test than a control group that had only received metalinguistic information. The group that had received both types of practice did better in production, but was slower. In the domain of phonology, Li and DeKeyser (in press) provided training in the perception and production of tones in Mandarin and obtained the same results as DeKeyser (1997): performance was far better when participants were tested on the skill practiced than when they were tested on the reverse skill, again for both reaction time and error rate.

Where ostensibly the same skill is required, but in a slightly different variant/context, the nature of the practice still determines whether transfer succeeds or not. What is needed is practice that leads to transfer-appropriate processing (TAP; Morris, Bransford, & Franks, 1977), that is, practice that has enough elements in common with the context of transfer for this context to activate the memory traces from the practice. This principle too is well known in cognitive and educational psychology (cf. Blaxton, 1989; Martin-Chang & Levy, 2005, 2006; Roediger, Gallo, & Geraci, 2002) and has support from cognitive neuroscience (Rugg, Johnson, Park, & Uncapher, 2008). Lightbown (2007) was perhaps the first to focus attention on this concept in applied linguistics. She drew on the work of various psychologists to argue that varying the conditions of practice is important to improve both depth of processing and transferability of learning. In particular, she posits that when L2 processing is entirely meaning-focused, that is not conducive to later retrieval of aspects of form. Spada, Jessop, Tomita, Suzuki, and Valeo (2014) provided some evidence for this point of view in the broad sense that learners in a group with isolated focus on form did better on an outcome test of written grammar, while those in a group that experienced focus on form integrated in communicative activities performed better on an oral production task.

Still more broadly speaking, the need for TAP can be seen as an example of the importance of depth of processing (Craik, 2002; Craik & Lockhart, 1972; Lockhart, 2002), which in turn explains to a large extent why making things more difficult during the learning process may lead to better results in the end (Linn & Bjork, 2006). Brown et al. (2014) stress this principle throughout their book and show how it applies to a wide variety of paradigms. A well-known example in cognitive psychology is Karpicke and Roediger's (2007, 2008) research on paired-associate vocabulary learning, which showed that the effect of repeated testing is stronger than the effect of repeated exposure. An example from SLA is Schneider, Healy, and Bourne (2002), who demonstrated that learners of French L2 vocabulary did better after more difficult English-to-French practice than after French-to-English and when there was no pretraining (as measured by reduced forgetting as well as enhanced savings during relearning).

Two issues that have drawn increasing attention in SLA, the ideal distribution of practice and the best way of providing CF, can also be seen against this background of enhanced learning through increased difficulty/depth of processing during practice. Much work in cognitive and educational psychology, in the laboratory (e.g., Pavlik & Anderson, 2005) as well as in the classroom (e.g., Sobel, Cepeda, & Kapler, 2011), has demonstrated the advantage of distributed over massed practice, in particular for vocabulary learning. Cepeda et al. (2006) meta-analyzed this literature and came to the conclusion that the ideal spacing depends on how delayed the testing is. Massed learning or learning at very short intervals may lead to good performance on immediate posttests, but not on delayed posttests, while long intervals between practice sessions may lead to somewhat lower scores on immediate posttests, but much higher scores on substantially delayed posttests. One way of interpreting this is that the harder the learning becomes, that is, the more memory is taxed because of the wide spacing, the more robust the learning is (for a thorough discussion of several alternative explanations, see Toppino & Gerbier, 2014). In SLA the findings have been less clear. Bird (2010) found distributed practice to be best for the learning of past tense use in ESL. In the same vein, Nakata (2015) showed an advantage for spacing in ESL vocabulary learning. Serrano and Muñoz (2007), however, found that for an English course as a whole, more concentrated teaching (25 hours of instruction per week during 5 weeks) was more effective than the same number of hours distributed over 3–4 months or 7 months. Suzuki and DeKeyser (in press-a, b) found no difference for different amounts of spacing, but did instead detect an aptitude-treatment interaction in the sense that the massed treatment drew more on participants' memory capacity and the distributed treatment more on analytic ability. Given the different focus and the different time scales of these studies, it is hard to pinpoint the reasons for the differences in their findings and to come to any generalizations at this point.

In research on CF, the work by Roy Lyster and his colleagues is especially relevant from the perspective of skill acquisition (for an overview, see Lyster & Sato, 2013). Sato and Lyster (2012), in particular, show how Japanese English as a Foreign Language (EFL) learners who had been trained to give each other CF did better in terms of both accuracy and fluency than those who had not, even though they received the same amount of interactional practice. Such dual improvement is hard to explain through notions such as noticing, but fits in well with a skill acquisition perspective, where both reaction time and error rate typically decline with continued practice. Interestingly, recasts and prompts yielded the same results. It should be pointed out that these learners entered the study with a large amount of declarative knowledge, but little proceduralization, let alone automatization, and it would be naïve to expect similar results with learners who do not have the prerequisite declarative knowledge: “there needs to be knowledge to be practiced” (DeKeyser, 2010b, p. 161). Only because of their considerable declarative knowledge were these learners able to use both recasts and prompts as reminders of what they already knew in principle, but what only became proceduralized through repeated cycles of feedback (or self-monitoring encouraged by both the feedback itself and the training for peer feedback they had received) and modified output. The results are particularly encouraging as the study was longitudinal (10 weeks) and carried out in a real classroom, and as the outcome measures tested fluency and accuracy overall, not just for the structures that were used in the training.

Similar interpretations, in the framework of skill acquisition theory, apply to the findings of Yang and Lyster (2010), with Chinese learners of EFL, focusing on past tense. In this study prompts worked better for regular forms and recasts for irregular forms, which the authors explain by saying that “prompts are predicted to help learners gain greater control

Robert DeKeyser

over already acquired forms and access them in faster ways, whereas recasts might be more effective for providing positive exemplars of relatively new linguistic forms” (p. 255). Here too, the learners had had ample opportunity to acquire declarative knowledge of the rules prior to the intervention, but needed form-focused practice in a communicative context to proceduralize their knowledge.

Teaching Tips

- *Stages of skill acquisition:* when planning activities, always think of whether they are meant to advance declarative knowledge, proceduralization, or automatization—is the aim to provide more understanding, getting to apply that understanding, or getting to use it faster, with less effort, more spontaneously?
- *Transfer-appropriate processing:* plan activities so that the mental processes learners go through are similar to later activities and eventually to activities in the real world, and make sure learners are conscious of how they have used their declarative knowledge in previous activities.
- *Require effort:* if learners have to make an effort to carry out a task, the relevant elements will be processed more deeply, and the knowledge retrieval processes are likely to be more similar to the ones needed in later, more realistic tasks.
- *Distribute practice, at least for declarative knowledge:* intense, focused practice followed by months of ignoring the same structures is less efficient than reminding learners of these structures each time memory begins to fade.
- *Provide corrective feedback, but in a way that is suited for what you want to achieve:* a recast may be better for teaching a new vocabulary item (declarative knowledge), but a prompt may be better for previously learned grammar (proceduralization), and corrective feedback may be of little use for grammar that was not thoroughly covered previously.
- *Individualize:* try to keep track of where individual students are with respect to a given structure: lack of understanding, near perfect understanding, ability to apply in easy contexts, ability to apply in new contexts, ability to apply under pressure . . . and adjust activities and feedback accordingly.

Future Directions

The previous sections have made it clear that the amount of recent research that is directly aimed at applying skill acquisition theory in SLA is rather limited, but that the number of studies that are relevant from this perspective is much larger. A first desideratum for future research, therefore, is more studies that test hypotheses that follow directly from the theory, for example, whether learning curves reflect a power function, or whether learners in advanced stages of proceduralization show evidence of automatization in the narrow sense of the word, not just speed-up.

Second, what is sorely needed from an applied perspective is studies that are longitudinal and are carried out in a classroom context, yet look closely at very specific processes in a controlled design, in other words, studies that combine ecological validity with internal validity. This is, of course, a tall order, and can probably only be achieved in a classroom context where computers are already used frequently, in order to allow for both the treatment and the outcome measures to be administered to the students in a context that is representative of the

classroom, yet allows for a strictly controlled treatment and a fine-grained documentation of the students' progress.

Meanwhile, several less demanding approaches can be very useful. In contexts where strict control over the treatment is impossible, we should at least strive for studies that look at the longitudinal development of a few structures in the classroom and beyond (for instance in study abroad contexts), but that take fine-grained measures, not just for accuracy, but also for fluency as well as frequency of spontaneous use. In contexts that are more constrained than the classroom or study abroad, yet representative of real-world learning processes, for instance regular practice with conversation partners, we may have a better opportunity to add an introspective component to the study and to combine a longitudinal perspective with a close look at individual differences.

Individual differences should also be a focus of research in computer-assisted learning contexts. Computers offer excellent opportunities for individualized practice, but more often than not at this point that is limited to individualization in terms of speed or number of items. Adaptation to individual aptitude profiles, proficiency profiles, or learning preferences can be accomplished, not only at the curricular level, but also on a minute-to-minute basis, using computer modeling of the student's skill acquisition profile (see, e.g., Koedinger, 2006; Nakic, Granic, & Glavinic, 2015). Collaborative research between language acquisition researchers, computational linguists, and educational technologists is sorely needed on this point if materials development for CALL is to become more sophisticated, in particular with respect to individualization of practice.

The same can be said for practice with conversation partners, which is now offered commercially by a few companies. How often should the sessions take place? How should they be linked to the students' other learning experiences? How should they be monitored and documented in order to allow for fine-tuning of future practice sessions? Most importantly, perhaps, how should the (paid) conversation partners be trained (in part based on the answers to the previous questions)?

Developments in information technology and communication infrastructure are providing ever more opportunities for communication in nonnative languages, but also for learning languages and for research on these learning processes. Skill acquisition theory is a better fit than most as a framework for such research and can inspire a variety of collaborative efforts between cognitive and educational psychologists, psycholinguists, neurolinguists, computational linguists, and SLA researchers.

References

- Bird, S. (2010). Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics*, 31(4), 635–650.
- Blaxton, T.A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 5(4), 657–668.
- Bowden, H. W., Gelfand, M. P., Sanz, C., & Ullman, M. T. (2010). Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language Learning*, 60(1), 44–87.
- Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, 51, 2492–2511.
- Brown, P. C., Roediger, H.L.I., & McDaniel, M.A. (2014). *Make it stick: The science of successful learning*. Cambridge, MA: Belknap/Harvard.

Robert DeKeyser

- Carpenter, S.K., Cepeda, N.J., Rohrer, D., Kang, S.H.K., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review*, 24(3), 369–378.
- Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380.
- Chein, J., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, 25(3), 607–623.
- Corkin, S. (2013). *Permanent present tense: The unforgettable life of the amnesic patient, H.M.* New York: Basic Books.
- Craik, F.I.M. (2002). Levels of processing: Past, present . . . and future? *Memory*, 10, 305–318.
- Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- de Jong, N. (2005). Can second language grammar be learned through listening? An experimental study. *Studies in Second Language Acquisition*, 27(2), 205–234.
- de Jong, N., & Perfetti, C.A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 62(2), 533–568.
- DeKeyser, R.M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19(2), 195–221.
- DeKeyser, R.M. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). New York: Cambridge University Press.
- DeKeyser, R.M. (2007). Study abroad as foreign language practice. In R.M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208–226). New York: Cambridge University Press.
- DeKeyser, R.M. (2009). Cognitive-psychological processes in second language learning. In M. Long & C. Doughty (Eds.), *Handbook of second language teaching* (pp. 119–138). Oxford: Wiley-Blackwell.
- DeKeyser, R.M. (2010a). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals*, 43(1), 80–92.
- DeKeyser, R.M. (2010b). Practice for second language learning: Don't throw out the baby with the bathwater. *International Journal of English Studies*, 10(1), 155–165.
- DeKeyser, R.M. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 94–112). London: Routledge.
- DeKeyser, R.M., & Criado-Sánchez, R. (2012). Automatization, skill acquisition, and practice in second language acquisition. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 323–331). Oxford: Wiley-Blackwell.
- DeKeyser, R.M., & Prieto Botana, G. (2015). The effectiveness of processing instruction in L2 grammar acquisition: A narrative review. *Applied Linguistics*, 36(3), 290–305.
- Doughty, C.J. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C.J. Doughty & M.H. Long (Eds.), *Handbook of second language acquisition* (pp. 256–310). Oxford: Blackwell.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305–352.
- Ferman, S., Olshtain, E., Schechtman, E., & Karni, A. (2009). The acquisition of a linguistic skill by adults: Procedural and declarative memory interact in the learning of an artificial morphological rule. *Journal of Neurolinguistics*, 22, 384–412.
- Gade, M., Druey, M.D., Souza, A.S., & Oberauer, K. (2014). Interference within and between declarative and procedural representations in working memory. *Journal of Memory and Language*, 76, 174–194.
- Golonka, E.M. (2006). Predictors revisited: Linguistic knowledge and metalinguistic awareness in second language gain in Russian. *Modern Language Journal*, 90(4), 496–505.
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). Amsterdam: John Benjamins.

- Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, 11, 523–532.
- Hill, N. M., & Schneider, W. (2006). Brain changes in the development of expertise: Neuroanatomical and neurophysiological evidence about skill-based adaptations. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 653–682). New York: Cambridge University Press.
- Hulstijn, J. H. (2002). Towards a unified account of the representation, acquisition, and automatization of second-language knowledge. *Second Language Research*, 18(3), 193–223.
- Hulstijn, J. H., van Gelderen, A., & Schoonen, R. (2009). Automatization in second language acquisition: What does the coefficient of variation tell us? *Applied Psycholinguistics*, 30(4), 555–582.
- Jiang, N. (2012). *Conducting reaction time research in second language studies*. London: Routledge.
- Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, 64(4), 809–854.
- Karpicke, J. D., & Roediger, H. L. I. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57, 151–162.
- Karpicke, J. D., & Roediger, H. L. I. (2008). The critical importance of retrieval for learning. *Science*, 319, 966–968.
- Kasper, G., & Kellerman, E. (1997). *Communication strategies: Psycholinguistic and sociolinguistic perspectives*. New York: Addison Wesley Longman.
- Kelly, A. M. C., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex*, 15(8), 1089–1102.
- Koedinger, K. R., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–77). New York: Cambridge University Press.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Englewood Cliffs, NJ: Prentice-Hall.
- Krashen, S. D. (1985). *The input hypothesis*. London/New York: Longman.
- Lafford, B. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition*, 26(2), 201–225.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365.
- Li, M., & DeKeyser, R. (in press). Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning. *Studies in Second Language Acquisition*.
- Lightbown, P. (2007). Transfer appropriate processing as a model for classroom second language acquisition. In Z. Han (Ed.), *Understanding second language process* (pp. 27–44). Clevedon: Multilingual Matters.
- Lim, H., & Godfroid, A. (2015). Automatization in second language sentence processing: A partial, conceptual replication of Hulstijn, Van Gelderen, and Schoonen's 2009 study. *Applied Psycholinguistics*, 36(5), 1247–1282.
- Linn, M. C., & Bjork, R. A. (2006). The science of learning and the learning of science: Introducing desirable difficulties. *APS Observer*, 19. Retrieved from <http://www.psychologicalscience.org/index.php/publications/observer/2006/march-06/the-science-of-learning-and-the-learning-of-science.html>
- Lockhart, R. S. (2002). Levels of processing, transfer-appropriate processing, and the concept of robust encoding. *Memory*, 10, 397–403.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32(2), 265–302.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40.
- Lyster, R., & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In M. García Mayo, J. Gutierrez-Mangado, & M. Martínez Adrián (Eds.), *Multiple perspectives on second language acquisition* (pp. 71–91). Amsterdam: John Benjamins.

Robert DeKeyser

- Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *Modern Language Journal*, 90(3), 320–337.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22(4), 471–499.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 407–452). New York: Oxford University Press.
- Martin-Chang, S. L., & Levy, B. A. (2005). Fluency transfer: Differential gains in reading speed and accuracy following isolated word and context training. *Reading and Writing*, 18, 343–376.
- Martin-Chang, S. L., & Levy, B. A. (2006). Word reading fluency: A transfer appropriate processing account of fluency transfer. *Reading and Writing*, 19, 517–542.
- McNamara, T. F. (1995). Modelling performance: Opening Pandora's box. *Applied Linguistics*, 16(2), 159–179.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132(2), 297–326.
- Morgan-Short, K. (2012a). Declarative memory and knowledge. In P. Robinson (Ed.), *Encyclopedia of second language acquisition* (pp. 157–160). London: Routledge.
- Morgan-Short, K. (2012b). Procedural memory and knowledge. In P. Robinson (Ed.), *Routledge encyclopedia of second language acquisition* (pp. 509–512). London: Routledge.
- Morgan-Short, K., Farett-Stutenberg, M., Bill-Schuetz, K. A., Carpenter, H., & Wong, P. C. M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, 17(1), 56–72.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, 60(1), 154–193.
- Morris, D. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519–533.
- Nakata, T. (2015). Effects of expanding and equal spacing on second language vocabulary learning: Does gradually increasing spacing increase vocabulary learning? *Studies in Second Language Acquisition*, 37(4), 677–711.
- Nakic, J., Granic, A., & Glavinic, V. (2015). Anatomy of student models in adaptive learning systems: A systematic literature review of individual differences from 2001 to 2013. *Journal of Educational Computing Research*, 51(4), 459–489.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29(4), 557–581.
- Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, 19, 1730–1737.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.
- Paulston, C. B. (1972). The sequencing of structural pattern drills. *TESOL Quarterly*, 6, 197–208.
- Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, 29, 559–586.
- Reber, A. S. (2003). Implicit learning. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (Vol. 2, pp. 486–491). London: Macmillan.
- Reber, P. J. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia*, 51, 2026–2042.

- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63(3), 595–626.
- Robinson, P., Mackey, A., Gass, S., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247–267). London: Routledge.
- Rodgers, D. M. (2011). The automatization of verbal morphology in instructed second language acquisition. *IRAL*, 49, 295–319.
- Roediger, H. L. I., Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, 10(5/6), 319–332.
- Rohrer, D. (2015). Student instruction should be distributed over long time periods. *Educational Psychology Review*, 2, 635–643.
- Rugg, M. D., Johnson, J. D., Park, H., & Uncapher, M. R. (2008). Encoding-retrieval overlap in human episodic memory: A functional neuroimaging perspective. In W. S. Sossin, J. C. Lacaille, V. F. Castelucci, & S. Belleville (Eds.), *Progress in brain research* (Vol. 169, pp. 339–352). Amsterdam: Elsevier.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Philadelphia/Amsterdam: John Benjamins.
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34(4), 591–626.
- Schneider, V. I., Healy, A. F., & Bourne, L. E., Jr. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, 46, 419–440.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. London: Routledge.
- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26(2), 173–199.
- Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, 14, 369–385.
- Segalowitz, S. J., Segalowitz, N. S., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics*, 19(1), 53–67.
- Serrano, R. (2011). The time factor in EFL classroom practice. *Language Learning*, 61(1), 117–145.
- Serrano, R., & Muñoz, C. (2007). Same hours, different time distribution: Any difference in EFL? *System*, 35, 305–321.
- Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning*, 63(2), 296–329.
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.
- Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real world classroom vocabulary learning. *Applied Cognitive Psychology*, 25(5), 763–767.
- Spada, N., Jessop, L., Tomita, Y., Suzuki, W., & Valeo, A. (2014). Isolated and integrated form-focused instruction: Effects on different types of L2 knowledge. *Language Teaching Research*, 18(4), 453–473.
- Spada, N., & Lightbown, P. (2012). Instructed second language acquisition. In P. Robinson (Ed.), *Routledge encyclopedia of second language acquisition*. London: Routledge.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308.
- Squire, L. R., & Wixted, J. T. (2011). The cognitive neuroscience of human memory since H.M. *Annual Review of Neuroscience*, 34, 259–288.
- Suzuki, Y., & DeKeyser, R. (2015). Does elicited imitation measure implicit knowledge? Evidence from the word-monitoring task. *Language Learning*, 65(4), 860–895.

Robert DeKeyser

- Suzuki, Y., & DeKeyser, R. (In press-a). Exploratory research on L2 distributed practice: An aptitude-by-treatment interaction. *Applied Psycholinguistics*.
- Suzuki, Y., & DeKeyser, R. (In press-b). Effects of distributed practice on the automatization of L2 morphosyntax. *Language Teaching Research*.
- Suzuki, Y., & DeKeyser, R. (submitted). The interface of explicit and implicit knowledge in a second language.
- Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, 16(2), 367–382.
- Thai, C., & Boers, F. (2016). Repeating a monologue under increasing time pressure: Effects on fluency, complexity, and accuracy. *TESOL Quarterly*, 50(2), 369–393.
- Toppino, T. C., & Gerbier, E. (2014). About practice: Repetition, spacing, and abstraction. *Psychology of Learning and Motivation*, 60, 113–189.
- Towell, R. (2012). Complexity, accuracy and fluency from the perspective of psycholinguistic second language acquisition research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 47–69). Amsterdam, The Netherlands: John Benjamins.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14(2), 103–135.
- Ullman, M. T. (2015). The declarative/procedural model. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 135–158). London: Routledge.
- Williams, J. (2009). Implicit learning in second language acquisition. In T. Bhatia & W. Ritchie (Eds.), *The new handbook of second language acquisition* (pp. 319–353). Bingley: Emerald.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235–263.

Intentional and Incidental L2 Learning

Ronald P. Leow and Celia C. Zamora

Background

Whether one can learn a foreign or second language (L2) incidentally, typically defined as the absence of any deliberate intention to learn target L2 information in the L2 input, has been an issue that has not only permeated several studies in the non-second language acquisition (SLA) field for decades but also in the current field of SLA. Indeed, the empirical origins of incidental learning date back to the beginning of the 20th century, and began in psychology-based studies (e.g., Jenkins, 1933). This chapter (1) succinctly traces the evolution of the study of incidental learning, as opposed to intentional learning, in relation to its definition, the target of investigation, the research methodology employed, and empirical findings, (2) discusses the roles of type of learning in light of current theoretical, methodological, and empirical issues within an SLA context, and (3) provides directions for future instructed SLA (ISLA) studies.

Intentional learning, defined recently as “a deliberate attempt to commit factual information to memory” (Hulstijn, 2013, p. 2632), or referred to as “cognitive processes that have learning as a goal rather than an incidental outcome” (Bereiter & Scardamalia, 1989, p. 363), has always been assumed to represent the type of learning, of a more explicit nature, that underscores a formal instructional classroom setting. The definition is relatively stable in many studies, albeit with some nuances as will be discussed herein, and clearly addresses some depth of processing or cognitive effort employed by the learner during the L2 learning process. However, a cursory review of what comprises incidental learning reveals quite a range of perceptions pertaining to what it actually entails and these are typically reflected in the methodology employed to address its role in the L2 learning process.

Non-SLA Field

One of the early and relatively broad definitions of incidental learning was provided by a psychologist, Jenkins (1933), who wrote that incidental learning is “learning which occurs in the absence of a specific intent to remember” (p. 471). This definition appears to refer to a relatively low level of processing or processing without much cognitive effort or subsequent

mental elaboration to retain the information in memory. The target items analyzed in early incidental learning studies were typically lexical-based and ranged, for example, from the learning of syllables (Jenkins, 1933), to pronunciation of trigrams (Mechanic, 1964), to associated/unassociated words (Hyde & Jenkins, 1973). The context for incidental learning later changed from simply lacking the intent to remember by the participant to the distinguishing features of instructional stimuli (e.g., the presence or absence of explicit instruction to learn), particularly those instructions that do not prepare the learners for retention of the material (e.g., Postman, 1964).

The inclusion of orienting tasks was first found in studies of incidental learning in the late 1960s/early 1970s. These orienting tasks were employed to further facilitate processing of the target stimuli under incidental conditions without mentioning the recall tasks or other forms of assessment of the target stimuli (e.g., Craik & Lockhart, 1972; Eysenck, 1982). According to Craik and Lockhart (1972), “the experimenter has a control over the processing the subject applies to the material that he does not have when the subject is merely instructed to learn and uses an unknown coding strategy” (p. 677). These orienting tasks could consist of tasks such as pleasant-unpleasant ratings, estimating the frequency of usage of the stimuli, sentence fragment judgment, and so forth (e.g., Hyde & Jenkins, 1973).

Two types of orienting tasks have been used in the incidental versus intentional learning paradigms: In the first type, participants performed an orienting task based on the stimulus materials, but were not given explicit instructions. In the second type, all participants, regardless of learning condition, were provided instructions to learn some of the stimuli; however, there were some additional stimuli included that participants were not explicitly told to process. The stimuli could be extrinsic (that is, including materials not part of the stimuli participants were instructed to learn) or intrinsic (additional components of the existing stimuli, for example, colors), and were the basis for the assessment of incidental learning (Eysenck, 1982).

From a theoretical perspective, early definitions of and studies on the role of type of learning (incidental versus intentional) in the learning process framed such learning in relation to the role of memory. To account for differential performances between intentional and incidental learning of vocabulary, Craik and Lockhart (1972) went a step further and postulated their levels of processing framework that focused on *how* learners processed the information in relation to memory. According to Craik and Lockhart, recalling information goes beyond having attended to it during its occurrence or having rehearsed it after its occurrence. Recollection depends also on how deeply this information was processed, namely, shallowly or deeply in relation to how much cognitive effort, elaboration rehearsal, and deeper analysis (such as activation of prior knowledge and meaningful analysis) was involved in the decoding of the incoming data. In a series of 10 experiments on word or lexical processing, Craik and Tulving (1975) reported overall empirical evidence for the effects of levels of processing on both incidental and intentional memory performance. It was assumed, then, that the explicit instructions to learn facilitate the learner’s processing of the material in a more effective manner than the incidental orienting task, which would account for the superiority of intentional over incidental learning (Craik & Lockhart, 1972; Postman, 1964). At the same time, if an appropriate incidental learning condition (such as a well-developed orienting task) were to facilitate deeper processing, and compared to an inferior intentional strategy, “learning under incidental conditions could be superior to that under intentional conditions” (Craik & Lockhart, 1972, p. 677). In other words, it appears that it may not be the experimental learning conditions that matter but *how* the target stimuli are processed by the learner.

In studies on L1 word lists, the notion of *transfer appropriateness* (Bransford, Franks, Morris, & Stein, 1979) has also been postulated to account for the typical superiority demonstrated by the intentional learning condition over the incidental one. This postulation is based on the compatibility between learning condition (e.g., read and pay attention to target words in the text) and testing measures (e.g., select words that you recognize from the text). For example, learners exposed to similar learning conditions and assessment tasks (e.g., +semantic/+semantic) were reported to have retained significantly more words when compared to those exposed to incompatible learning conditions and testing (e.g., +semantic/−semantic).

Two major assumptions in these aforementioned studies were that (1) type of experimental learning condition and instructions would lead to differential types of learning,¹ although learning was typically measured offline, and (2) intentional learning was superior based on postexposure tests (Eysenck, 1982).

Key Concepts

Incidental learning: Learning without any intention to learn.

Intentional learning: Learning with intent to learn.

Cognitive effort: The mental work involved in making decisions.

Orienting tasks: Specific instructions provided in a task to draw participants' attention to particular feature(s) in the stimuli.

SLA Field

The field of SLA has addressed the roles of intentional and incidental learning in the L2 learning process from quite a multifaceted perspective (Leow, 2015a). Theoretically, the notions of intentional and incidental learning, from both a vocabulary and grammatical perspective, appear to have a close connection to Krashen's (1982) Monitor Model that can be deconstructed from three perspectives (Leow & Cerezo, 2016). The first is that the acquisition *process* is subconscious, that is, without awareness. Awareness may be defined as “a particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus” (Tomlin & Villa, 1994, p. 193) and is typically associated with type of learning, namely, explicit learning (learning with awareness) and implicit learning (learning without awareness). The second perspective is *how* the L2 data are processed during acquisition. Acquisition is viewed as being “effortless” on the part of the learner who processes the language with minimal amount of cognitive effort. The third perspective regards the context in which acquisition occurs, namely, a language environment in which exposure to and interaction with the target language is prominent. Krashen (1982) also described acquisition as the following: “implicit learning, informal learning, and natural learning. In non-technical language, acquisition is picking up a language” (p. 10), which appears to indicate that acquisition, incidental learning, and implicit learning all share two important features, namely, a lack of cognitive effort and an absence of awareness during the learning process. This conflation is seen in this direct association between the acquisition process and incidental learning: “Thus, the acquisition process is identical to what had been termed ‘incidental learning’” (R. Ellis, 1994, p. 212).

Key Concepts

Awareness: “A particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus” (Tomlin & Villa, 1994, p. 193).

Explicit learning: Learning with awareness.

Implicit learning: Learning without awareness.

Empirically, rather than focusing on memory and cognition, the early studies in SLA on incidental versus intentional learning investigated the effects of type of learning condition (intentional vs. incidental) on vocabulary learning through reading (e.g., Hulstijn, 1989; Krashen, 1989). Typically, underlying the motivation for the studies was a reference to the tremendous amount of vocabulary knowledge exhibited by L1 learners who clearly could not have learned all lexical items within a formal instructional setting (see Grabe, 2009 for an overview). Such vocabulary was more likely to have been “picked up” (see Krashen, 1982), which soon became associated with the notion of incidental learning because the vocabulary was not intentionally learned or was not the primary focus of the learner.

The role of attention, signaling a shift to learner internal processes, also began to appear in definitions of incidental learning associated with the notion of “picking up” (R. Ellis, 1994; Schmidt, 1994) and also in the assumption that making a mental effort while reading had a positive effect on vocabulary learning (e.g., Hulstijn, 1992). The late ’90s witnessed a sharper focus on the roles of constructs such as attention and noticing (e.g., Robinson, 1997; Schmidt, 1990) in relation to grammatical items in the L2 data, rather than vocabulary. For example, Robinson (1997) framed incidental learning conditions as an exercise in understanding the meaning of discrete sentences that “replicates the learning condition that Krashen argues leads to unconscious acquisition (processing and understanding the meaning of input without intentionally focusing on grammatical form)” (p. 230). In addition, incidental learning was associated with implicit or unconscious learning that was postulated to be memory-based, item-specific, and nongeneralizable (e.g., Shanks & St. John, 1994) and lacking a focus on form when compared to enhanced or instructed conditions with a focus on form (Robinson, 1997).

Methodologically, the research designs employed in many of these incidental vocabulary and grammatical studies were relatively similar to those employed in the psychology-based studies to address the quantitative aspects or qualitative properties of incidental learning during exposure to a reading text or L2 grammatical data. The majority of the reading studies provided the participants with a reading comprehension task, where, in the intentional learning condition, the target words were included with a dictionary, gloss, contextual clues, or some manner with which the participant could infer the meaning. The grammatical studies typically included a training phase in which the stimuli comprised multiple exemplars of the target word, form, or structure.

Current Issues

Current SLA studies from the 2000s continue to address the following general theoretical question: does L2 learning of target information in the L2 input take place without any deliberate attempt to do so (incidental learning), that is, when the primary focus of the learner is on other features of the L2 input? One example may be processing for content

information versus processing for lexical or grammatical information. However, the concepts of intentional and incidental learning began to be associated with depth of processing, recently defined as “the relative amount of cognitive effort, level of analysis, elaboration of intake together with the usage of prior knowledge, hypothesis testing and rule formation employed in decoding and encoding some grammatical or lexical item in the input” (Leow, 2015a, p. 204), or conflated with explicit and implicit learning, that is, learning with or without awareness, respectively. Within these learning conditions, studies also began to address methodologically (1) the process, that is, *how* learners process the L2 data (e.g., the role of attention or awareness or lack thereof), and (2) the product, that is, type of learner knowledge (implicit versus explicit) as measured after the experimental learning exposure. This awareness of different stages along the L2 learning process (Leow, 2015a, 2015b) has led to the current methodological issue of operationalizing and measuring the construct of awareness, assumed to play an important role in differentiating type of learning. Studies have also sought to address other independent variables, for example, frequency of target items (Hamrick & Rebuschat, 2014) and individual differences (e.g., Grey, Williams, & Rebuschat, 2015; Kachinske, Osthus, Solovyeva, & Long, 2015; Robinson, 2005, 2010), within this learning condition strand of research.

Attention/Depth of Processing

Within the incidental L2 vocabulary learning strand, Godfroid, Boers, and Housen (2013) recently employed the concurrent procedure of eye-tracking to establish the role of attention in incidental L2 vocabulary learning (see Leow, Grey, Marijuan, & Moorman, 2014 for a critical discussion of concurrent data elicitation procedures in SLA). Laufer and Hulstijn's (2001) involvement load hypothesis was proposed to support the roles of attention (e.g., Schmidt, 1990) and cognitive or mental effort (e.g., Craik & Lockhart, 1972) deemed crucial for vocabulary retention. Several studies have tested Laufer and Hulstijn's involvement load hypothesis (e.g., Keating, 2008; Kim, 2008; Martínez-Fernández, 2008; Rott, 2005). Depth of processing has been associated with levels of awareness (Leow, 2012, 2015a) and postulated to play an important role in the intake processing stage along the L2 learning process (Leow, 2015a).

Key Concepts

Depth of processing: “The relative amount of cognitive effort, level of analysis, elaboration of intake together with the usage of prior knowledge, hypothesis testing and rule formation employed in decoding and encoding some grammatical or lexical item in the input” (Leow, 2015a, p. 204).

Incidental/Implicit Versus Intentional/Explicit

The role of awareness or lack thereof began to be addressed in several permutations of learning conditions. These learning conditions employed basically a similar type of incidental learning condition design in which all participants were provided with instructions to learn some of the experimental data, as discussed earlier. More specifically, some studies began to address the role of intentional or explicit learning in the L2 learning process in opposition

to the role of implicit learning. Implicit learning was defined as learning without awareness and with no intention to learn. However, this type of learning was typically operationalized and measured after the experimental exposure, as observed in definitions of implicit learning such as “the process that occurs when an item is learned without intention or awareness” (Kachinske et al., 2015, p. 387; see also Leung & Williams, 2011, 2012; Williams, 2005). Other studies probed deeper into type of resultant knowledge (implicit vs. explicit) exhibited after an incidental learning condition or exposure (e.g., Grey, Williams, & Rebuschat, 2014; Hamrick & Rebuschat, 2014; Rebuschat, Hamrick, Sachs, Riestenberg, & Ziegler, 2013; Rebuschat & Williams, 2012; Rogers, Révész, & Rebuschat, 2016). This opposition between implicit/incidental learning and explicit/intentional learning appears to be mainly derived from both Krashen’s (1982) Monitor Model and the field of cognitive psychology (see Reber’s seminal 1967 and other studies that investigated these types of learning employing artificial grammars or finite-state grammars that generate meaningless letter strings). At the same time, in an effort to establish the role of awareness or lack thereof during the L2 learning process, that is, *how* participants were processing the L2 data, other studies (e.g., Hama & Leow, 2010; Leow, 1997, 2000; Rosa & Leow, 2004; Rosa & O’Neill, 1999; Sachs & Suh, 2007) were employing a concurrent data elicitation procedure. This procedure elicited nonmetacognitive think aloud protocols, in which participants were requested to say aloud what they were thinking as they performed the experimental task without any explanation provided for their thoughts. Protocols were subsequently coded to establish empirically the presence or absence of awareness before any statistical analyses were performed to address its role in learning.

The empirical effort to address more directly learners’ cognitive processes, and more particularly, the construct of awareness, has led to the current methodological debate in the strand of research purporting to address its role in the L2 learning process (e.g., Hama & Leow, 2010; Leow, 2015a, 2015b; Leow & Hama, 2013; Leung & Williams, 2011, 2012; Rebuschat, Hamrick, Sachs, Riestenberg, & Ziegler, 2015), which is not difficult to extend to incidental and intentional learning condition studies. This debate has highlighted a crucial difference between stages at which cognitive constructs (e.g., attention, awareness) are being addressed. The first stage is at the concurrent (online) or construction stage of accessing and encoding the incoming experimental information. Operationalizing a cognitive construct at this stage views learning as a *process* and provides a richer insight into the actual point of encoding and decoding the L2. The second stage is at the nonconcurrent (offline) or reconstruction stage of retrieval of stored knowledge of the target linguistic rule or word and is viewed as a *product* (see Leow, Johnson, & Zárate-Sáñez, 2011 for further elaboration on stages and Leow, 2015a, 2015b for a distinction between the process of learning measured concurrently versus the product of learning measured nonconcurrently). As pointed out in Leow and Hama (2013), failure to gather concurrent data to establish that some cognitive construct did indeed play a role during the learning process may lead to an internal validity issue, that is, whether the findings faithfully reflect what the study set out to investigate.

Indeed, the typical research design of many of the studies purporting to address the roles of type of learning (e.g., incidental, intentional, implicit, explicit) *during* the learning process employed offline (awareness) measures administered after the experimental phase or treatment. However, some caution is warranted in the interpretation of the data gathered at this offline stage. Like the early intentional and incidental studies in psychology, there are minimally three major assumptions associated with this research design. The first is that all participants in either experimental condition are going to behave according to the assigned

condition. In other words, those in the intentional learning condition will make an elaborated attempt, that is, demonstrate cognitive effort, to learn the target information in the experimental data. Those in the incidental learning condition will process primarily one feature of the experimental data (e.g., data content) and simultaneously learn, pick up, or process the target information at a very low level. The second assumption is that performance on after-exposure awareness measures will reflect the learning behavior of each experimental learning condition. This operationalization of type of learning assumed to have taken place *during* an experimental learning condition does not control or guarantee what learners actually did in the condition. For example, some learners assigned to the incidental learning condition might have attempted to learn something in the input, that is, they might have entered the condition without any intention to learn something but during the exposure did intentionally try to learn something, especially after noticing mismatches between their L1 and the L2 input. Establishing what learners actually did during the experimental exposure does provide some confidence in the findings while making assumptions on internal processes may not be robust for scientific research. It is well established in the SLA field that, based on both concurrent (think aloud protocols) (e.g., Alanen, 1995; Hama & Leow, 2010; Leow, 1997, 1998a, 1998b, 2000; Rosa & Leow, 2004; Rosa & O'Neill, 1999) and nonconcurrent (post-exposure questionnaires) (e.g., Robinson, 1996, 1997) data, participants within experimental cells do not all behave according to assigned experimental condition. For example, Leow's (2000) think aloud protocols revealed that half of the participants processed deeply while the other half did not, notwithstanding being exposed to the same L2 data, and Rebuschat et al. (2015) revealed that participants in the same incidental learning condition demonstrated both implicit and explicit knowledge after exposure; see also Hamrick and Rebuschat (2014), Robinson (2002, 2005), Rogers et al. (2016).

The third assumption is that the amount of time participants are provided to process the experimental data is adequate to promote some kind of incidental or implicit learning. If one were to simulate a learning condition in which there is almost no depth of processing or minimal cognitive effort to learn new information, then participants need to be provided with a very short time span to eliminate potential deeper processing. A cursory survey of studies investigating implicit or incidental learning easily reveals a relatively large amount of time participants were provided to respond during exposure to the target data, ranging from about two seconds (e.g., Kachinske et al., 2015) to 20 seconds (Chen et al., 2011). In some studies, participants also performed a picture description task, a sentence reformulation task, and/or received feedback (e.g., Leung & Williams, 2011, 2012; Williams, 2005) that were *assumed* to distract participants from focusing on the target data. It is also of interest to note that replication (Martínez-Fernández, 2008) or extension (Hama & Leow, 2010) studies that have employed concurrent data elicitation procedures provide quite different results from the original studies that addressed vocabulary (Laufer & Hulstijn, 2001) and grammatical (Williams, 2005) learning, respectively. As can be seen, the use of experimental learning conditions to operationalize a learning process, be it incidental, intentional, implicit, or explicit, is not without internal validity limitations and may lead to a Type I or Type II error. A Type I or Type II error either over- or underestimates the effect of the learning condition (see Leow & Hama, 2013 for further elaboration). At the same time, it is commendable that some recent studies employing nonconcurrent data elicitation procedures have been more careful to report the effects of type of learning condition on type of knowledge (implicit vs. explicit) instead of attempting to extrapolate the findings to the process of learning, that is, at the encoding stage (e.g., Grey et al., 2014; Hamrick & Rebuschat, 2014; Rebuschat et al., 2013; Rebuschat & Williams, 2012; Rogers et al., 2016).

It is noteworthy that in the research design of the majority of current studies addressing type of learning or knowledge (e.g., Bordag, Kirschenbaum, Tschirner, & Opitz, 2014; Grey et al., 2014, 2015; Hamrick & Rebuschat, 2014; Kachinske et al., 2015; Leung & Williams, 2011, 2012, 2014; Rebuschat et al., 2013; Rebuschat & Williams, 2012; Rogers et al., 2016; Williams, 2005; Williams & Kuribara, 2008) there are two dominant features. The first is the relatively popular use of a (semi)artificial language or lexicon as the experimental L2 input (see Robinson, 2010 for a critique in relation to extrapolating findings from artificial grammar (AG) studies to naturally occurring languages due to a failure to find correlations between his AG group and the naturally occurring Samoan language group). The second is the use of nonconcurrent or offline measures (see Leow & Hama, 2013 for a critique in relation to addressing the process of learning) that may include almost exclusively grammatical or acceptability judgment tasks, offline verbal reports, and subjective awareness measures such as confidence level and source attributions (both self-reports) (see Rebuschat, 2013 for further elaboration of these offline awareness measures).

In sum, there appears to be some conflation between incidental learning (typically associated with “picking up” a language and opposed to intentional learning) and implicit learning (typically associated with a lack of awareness and opposed to explicit learning). There is also a current methodological debate that has highlighted a crucial difference between stages (concurrent/construction vs. nonconcurrent/reconstruction) at which cognitive constructs are being addressed. In addition, it has been recommended that incidental and implicit learning condition studies employing semi-artificial experimental data and after-exposure awareness measures exercise some caution in data interpretation when extrapolating their findings to naturally occurring languages.

Empirical Evidence

The outcomes of intentional and incidental learning have been measured by a variety of instruments in SLA studies. For vocabulary, these include, for example, the Vocabulary Knowledge Scale (VKS, Wesche & Paribakht, 1996), multiple-choice, retention, recognition, recall, vocabulary comprehension, lexical decision, semantic priming tests, self-paced reading, and so on.

Studies addressing grammatical learning or knowledge included instruments such as picture-matching tests, acceptability or grammaticality judgment tasks, morphological and syntactic tests, and reaction times while several of these judgment tests, together with offline verbal reports, confidence ratings, and source attributions were also employed to measure the construct of awareness or lack thereof. The popular cognitive psychology-based statistical analysis, namely, the chance test, has also been employed in some of these studies (e.g., Hama & Leow, 2010; Hamrick & Rebuschat, 2014; Williams, 2005). In a chance test, any mean score statistically above chance (50%) was reported as evidence of learning having taken place (see Williams, 2005).

Quite a range of target items has also been empirically investigated and these include artificial determiners encoding distance and animacy, pseudoclefts of location in English, word order, morphosyntax, dative alternation, semi-artificial languages and pseudo words, non-native syntax such as Japlish (sentences with Japanese syntax and case markers but English lexis) and Japanese scrambling (an optional syntactic operation that moves a phrase in the direction opposite to the head direction) in word order, locative markers, and case markings.

Text length for vocabulary studies ranged from short simplified texts of 100 words to novels consisting of 67,000 words. Data sets (including both lexical and grammatical

items) ranged from 8 to over 380 exemplars, with a range of untimed or approximately 2–20 seconds of exposure to each individual exemplar in a data set and an overall exposure from 10 minutes to 13 weeks. Training phases lasted from 10 minutes to over a number of days. Only a few studies administered delayed posttests (e.g., Grey et al., 2014, 2015; Robinson, 2002, 2005). Different levels of language experience have also been explored, ranging from no prior instruction on or knowledge of the target language or artificial data to advanced levels.

Even though some vocabulary learning has been reported in studies addressing reading naturally occurring languages (e.g., Barcroft, 2009; Hulstijn, 1992; see Ramos & Dario, 2015 for a critique) or pseudowords (presented in several exposure trials) in relation to frequency effects (Hamrick & Rebuschat, 2014), embedded within a naturally occurring language (Godfroid et al., 2013) or in relation to syntactic complexity (Bordag et al., 2014; Rogers et al., 2016), the robustness of learning leaves much to be desired. With regard to the involvement load hypothesis (Hulstijn & Laufer, 2001; Keating, 2008; Kim, 2008; Martínez-Fernández, 2008; Rott, 2005), several studies have provided empirical support, although Martínez-Fernández (2008) failed to do so after addressing several methodological limitations in the previous research designs, including the failure to use process measures to establish depth of processing.

Other studies have addressed incidental learning conditions or exposure on subsequent grammatical development of mostly artificial items embedded within naturally occurring phrases or sentences. They have also reported evidence that adults can learn aspects of nonnative syntax or morphosyntax while processing the language input for meaning and without any instruction to search for or learn a rule (e.g., Grey et al., 2014; Hamrick, 2014; Kachinske et al., 2015; Rebuschat & Williams, 2012; Robinson, 1995; Rogers et al., 2016; Williams & Kuribara, 2008). This evidence was based primarily on the results of the typical chance test, and was said to occur even after a delay of 2 weeks (e.g., Grey et al., 2014). This type of incidental learning condition can also lead to both implicit and explicit knowledge (e.g., Hamrick & Rebuschat, 2014; Rebuschat et al., 2013, 2015; Rebuschat & Williams, 2012; Rogers et al., 2016), as measured on grammaticality judgment tests. Some studies sought to explain such incidental learning, for example, of word order, in terms of associative (sequence) learning (e.g., Williams, 2010); some have relied on the role of awareness or lack thereof gleaned from awareness measures administered after the experimental exposure (e.g., Hamrick & Rebuschat, 2014; Leung & Williams, 2011, 2012; Rogers et al., 2016; Williams, 2005), while Leung and Williams (2014) addressed the role of prior knowledge in implicit learning and Kachinske et al. (2015) reported partial evidence for statistical learning.² However, like the vocabulary studies, the amount of learning reported after exposure, albeit relatively short in duration, is usually not robust.

At the same time, studies comparing intentional versus incidental learning conditions have typically reported that intentional learning conditions often result in more learning when compared to incidental learning conditions (e.g., Hamrick & Rebuschat, 2014). Similarly, studies that compared aware versus unaware learners (operationalized and measured either concurrently or nonconcurrently) also reported similar superior performance by the explicit learning group (e.g., Kachinske et al., 2015; Leow, 2000; Leung & Williams, 2011, 2012; Rebuschat & Williams, 2012, Experiment 1; Rebuschat et al., 2013). Mean percentages obtained by the unaware groups on the chance tests usually fell between a range of 49–61%, while the aware groups were substantially above this range, falling in the 70–88% range.

Future Directions

In spite of the relatively large number of studies that have empirically investigated type of learning (intentional and incidental together with implicit and explicit learning), there still remain theoretical, methodological, and pedagogical issues to be addressed in the SLA literature. First of all, what specifically does it mean to learn intentionally or incidentally? The broad definition of being requested to focus on some particular aspect of incoming L2 data (intentional) or picking up secondary information while the learner's primary attention is on another feature of the L2 data (incidental) is quite vague with regard to *how* specifically L2 learners process during incidental or intentional learning conditions. For example, does learning intentionally equate to learning explicitly, that is, with awareness, or does it mean that more cognitive effort is made but no guarantee that awareness of the target information or even learning are indeed achieved? Does learning incidentally equate to implicit learning or picking up the language, that is, learning without awareness or without any measurable amount of cognitive effort? Or does it mean that no intention to learn was present, but during exposure learners may become aware of target information or process deeply such information but with no guarantee that the target information is indeed learned? Is it possible that the mere fact that participants are being exposed to experimental materials, as in an empirical study, may raise some awareness of something to be learned and potentially tested afterwards (in spite of not being provided this priming or testing information)? It is the authors' perspective that learners do not enter experimental conditions without at least minimal intention to learn something. This perspective finds empirical support in several studies (e.g., Hamrick, 2014; Rebuschat et al., 2013, 2015; Rogers et al., 2016) that have reported intentional or explicit learning during a so-called incidental learning condition, defined as not providing specific instructions to learn any specific information in the L2 data or information regarding a postexposure test. What these basic questions reveal is a major concern of studies purporting to address the internal processes of adult L2 learners: the inadequacy of the operationalizations of these learning conditions in relation to assumed cognitive processes if concurrent data are not provided to support assumptions made on pre-exposure instructions and/or postexposure measures.

More specifically, how do studies addressing the roles of intentional and incidental learning relate to ISLA? This was defined recently as

a theoretically and empirically based field of academic inquiry that aims to understand *how* the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one's own.

Loewen, 2015, p. 2, emphasis added

What appear to be underscoring this definition (and others, see, for example, Housen & Pierrard, 2005) are (1) the focus on the "mechanisms of learning" (cognitive processes) employed in an instructed setting, that is, *how* L2 learners process L2 data in this setting as opposed to a more naturalistic setting; and (2) whether such processes can be manipulated by instructional intervention with the assumption that superior or faster L2 development will result. It may be instructive to situate future ISLA directions in relation to, for example, (1) a clearer definition of the construct of learning; (2) the operationalization of what constitutes type of learning (intentional, incidental, explicit, implicit), that is, *how*

L2 learners actually process the L2 data; (3) a more robust methodology to address type of learning; and (4) the context in which the learning is assumed to take place and its pedagogical implications.

First of all, future studies may want to address more precisely the construct of learning in (I)SLA. A cursory survey of published studies in both SLA and non-SLA fields reveals an inevitable mention of the term “learning.” At the same time, as pointed out in Leow (2015a), it is also revealing that what comprises “learning” within and between the SLA and non-SLA fields may not be the same construct. For example, the concept of intake (Corder, 1967) is postulated in many SLA theoretical models (e.g., Gass, 1997; Leow, 2015a; VanPatten, 2007) as information taken in during a preliminary stage along the L2 learning process but does not represent what is internalized in the L2 learner system. However, this concept is not well acknowledged in many non-SLA fields and whatever is taken in may be viewed as learning. Indeed, recent publications have discussed the construct of learning in reference to the role of memory (see Hulstijn, 2013) and memory traces (see Bordag et al., 2014), which may be associated with working memory from which initial data without further processing may disappear and not enter the L2 learner’s internal system. Other studies have viewed learning as a product that has been processed and eventually resides in the L2 learner’s internal system (e.g., Leung & Williams, 2012; Williams, 2005). In addition, there may be quite a lot of terminological confusion given that the construct of learning appears to be operationalized or measured by quite a wide range of assessment tasks, from simple recognition to production to grammaticality judgment tasks (see Leow, 2015a for a tri-dimensional perspective of the construct of learning in SLA).

To address type of learning, it may be important to revisit two key terms typically conflated in both the SLA and non-SLA literatures, namely, acquisition and learning (Leow, 2015a; Leow & Cerezo, 2016). The key distinctions between acquisition and learning lie precisely in *how* L1 and L2 learners process the L1 and L2 data, respectively (e.g., depth of processing, level of awareness, cognitive effort) and *where* exposure to the L1 and L2 occurs. In addition, the amount of time (and as an extension, the amount of target data) learners are exposed to and interacting with either the L1 or L2 needs to be seriously considered. In other words, viewed from this processing perspective and the ISLA formal and instructed context, two major distinctions between acquisition and learning are clearly based on type of processing (incidental/implicit vs. intentional/explicit, respectively) and type of context (naturalistic vs. instructed environment, respectively). More specifically, the typical ISLA formal setting, situated importantly within a language curriculum with its outcome goals, textbook, syllabi, limited exposure, tests, and so forth, is designed to promote more explicit and intentional learning than implicit and incidental learning and acquisition (see Leow & Cerezo, 2016, for a curricular approach to ISLA). This setting does not negate any instance(s) of incidental or implicit learning taking place in the formal instructed environment but, as Leow (2015a) cautions,

this kind of processing depends heavily on many factors that include the provision of large amounts of exemplars in meaningful contexts and quite a long period of time to process, internalize the exemplars, and have the knowledge available for subsequent usage.

p. 244

To address methodologically the process or mechanisms of learning, future research may want to make every effort to employ some concurrent data elicitation procedure

(e.g., eye-tracking, think aloud protocols) in the research design, if feasible, to gather data on learner processing and processes being employed while they are exposed to or interacting with the L2 data. The richness of concurrent data cannot be minimized and can certainly shed more light on internal processes or be used to triangulate data gathered at both online and offline stages (Leow, 2013; Rebuschat et al., 2015; Winke, 2013). Indeed, studies that have employed think aloud protocols have revealed robust L2 development associated with great depths of processing, high levels of awareness (hypothesis testing and rule formulation), and activation of both recently learned and prior knowledge (e.g., Cerezo, Caras, & Leow, 2016; Hsieh, Moreno, & Leow, 2015; Leow, 1997, 1998a, 1998b, 2000; Medina, 2015; Rosa & Leow, 2004; Rosa & O'Neill, 1999). Without concurrent data or empirical evidence to demonstrate that no intent or conscious effort was made during exposure to learn target items in the input (whether learners did, for example, pause at some target items and processed them with some level of cognitive effort or awareness or developed some strategy to process the L2 data), type of learning remains an unanswered question and, as noted earlier, ultimately lowers the level of internal validity of the study.

From a contextual perspective, it is not uncommon for researchers to premise their studies within an L1 acquisition perspective, that is, several references are made to the processes employed by L1 children with some assumption that their studies are being situated within a similar context. For example, with regard to the “picking up” of vocabulary, it is not controversial to note that the depth of processing exhibited by children acquiring their L1 may be relatively low and almost effortless. A similar contextual issue is found in incidental or implicit grammatical learning condition studies that appear to ground their theoretical underpinnings in child acquisition, for example, statistical learning (Saffran, 2003), sequence learning (Williams, 2010), or Krashen’s (1982) Monitor Model. Exposing L2 learners to an experimental written text or a series of data sets (oral or written) for less than an hour and then assuming that they will “pick up” (and, given the absence of delayed posttests, retain?) new vocabulary or grammatical information, even if presented multiple times, does not appear to acknowledge the following: (1) the huge disparity between L1 acquisition and L2 learning in regard to amount and type of exposure to and interaction with the L1 or L2 data, and (2) the depth of processing associated with type of learning. In addition, if pedagogical implications can be offered from studies investigating incidental learning conditions, researchers may need to address naturally occurring languages instead of the typical semi-artificial languages or lexicons employed in the research designs.

Probing deeper into the roles of incidental/implicit learning in adult L2 learning is of clear theoretical value to the field of SLA. However, viewed from both processing and contextual perspectives together with the empirical findings of demonstrated superiority of intentional and explicit learning over incidental and implicit learning, ISLA may better inform language curricula and teaching methodology by focusing on the potential roles either intentional or explicit learning (see also N. Ellis, 2015; Leow, 2015a) may play in promoting more robust learning in this setting. To this end, a strong ISLA research agenda may be to continue probing deeper into the cognitive processes employed by L2 learners as they interact with or are exposed to the L2 across different modalities, types of tasks, linguistic items, language levels, or instructions. A better understanding of these processes can contribute to the creation of theoretically based and empirically supported pedagogical tasks or activities that are designed to promote robust use of students’ mechanism of learning while performing such tasks or activities. This direction falls neatly within recent definitions of ISLA (e.g., Loewen, 2015).

Conclusion

This chapter has provided a succinct overview of the roles of intentional and incidental learning from its non-SLA root to current studies in the (I)SLA field. It has revealed the subtle changes in the definitions of what comprises both types of learning over the years, in the target of investigation, and in the research methodology employed. A critical discussion of these roles has also been provided in relation to current theoretical, methodological, and empirical issues within an SLA context, and, keeping closely to current definitions of ISLA, several directions for future ISLA studies are proposed.

Notes

1. This assumption is exemplified in Perruchet and Pacteau's (1991) statement: "That implicit learning follows from incidental instructions is a tacit assumption" (p. 4).
2. Statistical learning refers to one's ability to make use of statistical information in the input to support language acquisition. Early studies focused primarily on child acquisition.

References

- Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning and teaching* (pp. 395–411). Honolulu: University of Hawai'i Press.
- Barcroft, J. (2009). Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly*, 43(1), 79–103.
- Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 361–392). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bordag, D., Kirschenbaum, A., Tschirner, E., & Opitz, A. (2014). Incidental acquisition of new words during reading in L2: Inference of meaning and its integration in the L2 mental lexicon. *Bilingualism: Language and Cognition*, 18(3), 372–390.
- Bransford, J. D., Franks, J. J., Morris, C. D., & Stein, B. S. (1979). Some general constraints on learning and memory research. In L. S. Cerman & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 331–354). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cerezo, L., Caras, A., & Leow, R. P. (2016). Effectiveness of guided induction versus deductive instruction on the development of complex Spanish "gustar" structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, 38, 265–291.
- Chen, W., Guo, X., Tang, J., Zhu, L., Yang, Z., & Dienes, Z. (2011). Unconscious structural knowledge of form–meaning connections. *Consciousness and Cognition*, 20, 1751–1760.
- Corder, S. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161–170.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104(3), 268–294.
- Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). Philadelphia, PA: John Benjamins.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Eysenck, M. W. (1982). Incidental learning and orienting tasks. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 197–228). New York: Academic Press.
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.

Ronald P. Leow and Celia C. Zamora

- Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye tracking. *Studies in Second Language Acquisition*, 35(3), 483–517.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.
- Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*, 36(4), 611–645.
- Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences*, 38, 44–53.
- Hama, M., & Leow, R. P. (2010). Learning without awareness revisited. *Studies in Second Language Acquisition*, 32(3), 465–491.
- Hamrick, P. (2014). Recognition memory for novel syntactic structures. *Canadian Journal of Experimental Psychology*, 68, 2–7.
- Hamrick, P., & Rebuschat, P. (2014). Frequency effects, learning conditions, and the development of implicit and explicit lexical knowledge. In J. Connor-Linton & L. Amoroso (Eds.), *Measured language: Quantitative approaches to acquisition, assessment, processing, and variation* (pp. 125–139). Washington, DC: Georgetown University Press.
- Housen, A., & Pierrard, M. (2005). *Investigations in instructed second language acquisition*. Berlin: Mouton de Gruyter.
- Hsieh, H.-C., Moreno, N., & Leow, R. P. (2015). Awareness, type of medium, and L2 development: Revisiting Hsieh (2008). In R. P. Leow, L. Cerezo, & M. Baralt (Eds.), *A psycholinguistic approach to technology and L2 learning* (pp. 131–150). Berlin: Mouton de Gruyter.
- Hulstijn, J. H. (1989). Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input. In H. W. Dechert & M. Raupach (Eds.), *Interlingual processing* (pp. 49–73). Tübingen: Gunter Narr.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). London: Macmillan.
- Hulstijn, J. H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2632–2637). New York: Wiley-Blackwell.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539–558.
- Hyde, T. S., & Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, 12, 471–480.
- Jenkins, J. G. (1933). Instruction as a factor in “incidental” learning. *American Journal of Psychology*, 45(3), 471–477.
- Kachinske, I., Osthus, P., Solovyeva, K., & Long, M. (2015). Implicit learning of an L2 morphosyntactic rule, and its relevance for language teaching: Second language acquisition program. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 387–415). Philadelphia, PA: John Benjamins.
- Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, 12(3), 365–386.
- Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 58(2), 285–325.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal*, 73, 440–464.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26.
- Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, 47(3), 467–505.

- Leow, R. P. (1998a). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, 19, 133–160.
- Leow, R. P. (1998b). The effects of amount and type of exposure on adult learners' L2 development in SLA. *Modern Language Journal*, 82(1), 49–68.
- Leow, R. P. (2000). A study of the role of awareness in foreign language behavior. *Studies in Second Language Acquisition*, 22(4), 557–584.
- Leow, R. P. (2012). Explicit and implicit learning in the L2 classroom: What does the research suggest? *European Journal of Applied Linguistics and TEFL*, 2, 117–129.
- Leow, R. P. (2013). Schmidt's noticing hypothesis: More than two decades after. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor in Richard Schmidt* (pp. 23–35). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Leow, R. P. (2015a). *Explicit learning in the L2 classroom: A student-centered approach*. New York: Routledge.
- Leow, R. P. (2015b). Implicit learning in SLA: Of processes and products. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 47–65). Philadelphia, PA: John Benjamins.
- Leow, R. P., & Cerezo, L. (2016). Deconstructing the “I” and “SLA” in ISLA: One curricular approach. *Studies in Second Language Learning and Teaching*, 6, 43–63.
- Leow, R. P., Grey, S., Marijuan, S., & Moorman, C. (2014). Concurrent data elicitation procedures, processes, and the early stages of L2 learning: A critical overview. *Second Language Research*, 30(2), 111–127.
- Leow, R. P., & Hama, M. (2013). Implicit learning in SLA and the issue of internal validity. *Studies in Second Language Acquisition*, 35(3), 545–557.
- Leow, R. P., Johnson, E., & Zárate-Sáñez, G. (2011). Getting a grip on the slippery construct of awareness: Toward a finer-grained methodological perspective. In C. Sanz & R. P. Leow (Eds.), *Implicit and explicit conditions, processes and knowledge in SLA and bilingualism* (pp. 61–72). Washington, DC: Georgetown University Press.
- Leung, J. H. C., & Williams, J. N. (2011). The implicit learning of mappings between forms and contextually-derived meanings. *Studies in Second Language Acquisition*, 33(1), 33–55.
- Leung, J. H. C., & Williams, J. N. (2012). Constraints on implicit learning of grammatical form–meaning connections. *Language Learning*, 62(2), 634–662.
- Leung, J. H. C., & Williams, J. N. (2014). Crosslinguistic differences in implicit language learning. *Studies in Second Language Acquisition*, 29, 1–23.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. London: Routledge.
- Martínez-Fernández, A. (2008). Revisiting the involvement load hypothesis: Awareness, type of task and type of item. In M. Bowles, R. Foote, S. Perpiñán, & R. Bhatt (Eds.), *Selected proceedings of the 2007 second language research forum* (pp. 210–228). Somerville, MA: Cascadia Proceedings Project.
- Mechanic, A. (1964). The responses involved in the rote learning of verbal materials. *Journal of Verbal Learning and Verbal Behavior*, 3, 30–36.
- Medina, A. (2015). The variable effects of level of awareness and CALL versus non-CALL textual modification on adult L2 readers' input comprehension and learning. In R. P. Leow, L. Cerezo, & M. Baralt (Eds.), *A psycholinguistic approach to technology and L2 learning* (pp. 109–130). Berlin: Mouton de Gruyter.
- Perruchet, P., & Pacteau, C. (1991). Implicit acquisition of abstract knowledge about artificial grammar: Some methodological and conceptual issues. *Journal of Experimental Psychology: General*, 120, 112–116.
- Postman, L. (1964). Short-term memory and incidental learning. *Categories of Human Learning*, 24, 145–201.
- Ramos, R., & Dario, F. (2015). Incidental vocabulary learning in second language acquisition: A literature review. *Profile Issues in Teachers Professional Development*, 17(1), 157–166.
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 855–863.

Ronald P. Leow and Celia C. Zamora

- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research: A review. *Language Learning*, 63, 595–626.
- Rebuschat, P., Hamrick, P., Sachs, R., Riesterberg, K., & Ziegler, N. (2013). Implicit and explicit knowledge of form–meaning connections: Evidence from subjective measures of awareness. In J. Bergsleithner, S. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 249–270). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Rebuschat, P., Hamrick, P., Sachs, R., Riesterberg, K., & Ziegler, N. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition*, 37, 299–334.
- Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, 33, 829–856.
- Robinson, P. (1995). Aptitude, awareness, and the fundamental similarity of implicit and explicit second language learning. *Attention and Awareness in Foreign Language Learning*, 9, 303–357.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18(1), 27–67.
- Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, 19(2), 223–247.
- Robinson, P. (2002). Learning conditions, aptitude complexes, and SLA. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211–266). Philadelphia, PA: John Benjamins.
- Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, 27(2), 235–268.
- Robinson, P. (2010). Implicit artificial grammar and incidental natural second language learning: How comparable are they? *Language Learning*, 60, 245–263.
- Rogers, J., Révész, A., & Rebuschat, P. (2016). Implicit and explicit knowledge of inflectional morphology. *Applied Psycholinguistics*, 37, 781–812.
- Rosa, E. M., & Leow, R. P. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics*, 25(2), 269–292.
- Rosa, E. M., & O'Neill, M. D. (1999). Explicitness, intake and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511–556.
- Rott, S. (2005). Processing glosses: A qualitative exploration of how form–meaning connections are established and strengthened. *Reading in a Foreign Language*, 17(2), 95–124.
- Sachs, R., & Suh, B.-R. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second-language acquisition: A series of empirical studies* (pp. 197–227). Oxford: Oxford University Press.
- Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12, 110–114.
- Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Schmidt, R. W. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–210). London: Academic Press.
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human systems. *Behavioral and Brain Sciences*, 17, 367–447.
- Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183–203.
- VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 115–135). Mahwah, NJ: Lawrence Erlbaum.

- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, 53(1), 13–40.
- Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, 27, 269–304.
- Williams, J. N. (2010). Initial incidental acquisition of word order regularities: Is it just sequence learning? *Language Learning*, 60(2), 221–244.
- Williams, J. N., & Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. *Lingua*, 118, 522–553.
- Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension. *Studies in Second Language Acquisition*, 35(2), 323–352.

Complexity, Accuracy, and Fluency in L2 Production

Marije Michel

Background

Measuring the product of second language (L2) performance, that is, oral or written language, is a crucial aspect of research into instructed second language acquisition (ISLA) and has a long tradition. The earliest attempts to gauge performance in modern SLA research emerged in the 1970s and can be divided into two main strands (see Housen, Kuiken, & Vedder, 2012; Wolfe-Quintero, Inagaki, & Kim, 1998). First, based on research into first language (L1) acquisition, where mean length of utterance (MLU) was an established index of development, L2 researchers aimed for an index that would allow measurement of global L2 proficiency in reliable and valid ways and that would permit comparability over different studies and languages (see Larsen-Freeman, 1978). Second, from a pedagogical perspective, more and more classroom-based research into L2 performance started to characterize language use in terms of accuracy on the one hand and fluency on the other hand (Brumfit, 1979). Skehan (1998) added complexity and thereby introduced the triad of complexity, accuracy, and fluency (CAF) as the three fundamental dimensions characterizing L2 usage (Housen & Kuiken, 2009).

To date, the early working definitions of CAF are still used for global proficiency: Complexity refers to the size, elaborateness, richness, and diversity of the L2 performance. Accuracy is a measure for the target-like and error-free use of language. Fluency refers to the smooth, easy, and eloquent production of speech with limited numbers of pauses, hesitations, or reformulations. In the past two decades a growing body of research into ISLA has used CAF measures as dependent variables to gauge L2 performance manipulated by independent variables such as task complexity and task repetition. To a lesser extent some developmental studies have used CAF to identify change in quasi-experimental studies with pretest/posttest designs while others showcase longitudinal learner trajectories (for recent reviews see Housen & Kuiken, 2009; Housen et al., 2012; Lambert & Kormos, 2014 on CAF in general; Bulté & Housen, 2012, on complexity; Polio & Shea, 2014 on accuracy; Bosker, Pinget, Quené, Sanders, & de Jong, 2013 on fluency).

With respect to ISLA, Norris and Ortega (2009) state that

the primary reason for measuring L2 CAF is to account for how and why language competencies develop for specific learners and target languages, in response to particular tasks, teaching, and other stimuli, and mapped against the details of developmental rate, route, and ultimate outcomes.

p. 557

CAF dimensions are thought to be able to characterize different levels of L2 performance (Wolfe-Quintero et al., 1998). Furthermore, it is often assumed—although this is not always the case (see Lambert & Kormos, 2014; Pallotti, 2009)—that in comparison to less proficient L2 users or to themselves at earlier stages of development (e.g., before an instructional intervention), more proficient L2 learners (or after an instructional intervention):

1. Use a wider range of and more *complex* grammatical structures and vocabulary;
2. Produce more error-free utterances, that is, they are more *accurate*; and
3. Speak and/or write more *fluently*, that is, faster and with fewer instances of silence and repair.

In terms of cognitive processing, greater complexity and accuracy have been associated with a more elaborate and sophisticated L2 knowledge system related to representation and restructuring (or development) of the interlanguage while greater fluency is linked to more control and automatization, that is, faster access to L2 knowledge (Housen et al., 2012; Skehan, 2009).

Key Concepts

Complexity: Size, elaborateness, richness, and diversity of the learner's linguistic L2 system (Housen & Kuiken, 2009).

Accuracy: Degree of deviancy from a particular norm; deviations are usually characterized as errors (Wolfe-Quintero et al., 1998).

Fluency: Ease, eloquence, and smoothness of speech or writing (Chambers, 1997; Freed, 2000; Koponen & Riggensbach, 2000; Lennon, 1990).

The aim of this chapter is to give an overview of the CAF triad. In the next section, each of the three dimensions is presented with a definition, followed by a review of the challenges faced by current research. A final paragraph discusses ways to measure CAF. The following section will review empirical work that employed CAF and that provided experimental evidence relevant to ISLA. The next section will shed light on future directions in CAF research such as the role of communicative adequacy, the value of CAF when gauging interactive performance, and the use of advanced statistical methods and computer-based tools for CAF measurement. Finally, this chapter discusses the need for, on the one hand, standardization to increase validity, reliability, and generalizability of empirical work using CAF, and on the other hand, the need for (new) measures that are able to characterize the dynamic and organic system of L2 production and development.

Current Issues

With growing interest to use CAF as dependent variables to measure effects of manipulations on independent variables such as planning time, researchers have started to investigate the constructs more closely, with questions like: What exactly are we evaluating when measuring complexity? What is the ‘best’ measure to gauge accuracy? What are components of fluency? How do complexity, accuracy, and fluency and their subcomponents interact? Based on these reflections, the early assumption “that these three characteristics of language progress in tandem” (Wolfe-Quintero et al., 1998, p. 4) has now made room for the acknowledgment that complexity, accuracy and fluency are multifaceted, multi-layered and multidimensional in nature and that they are interrelated in complex and not necessarily linear ways (Housen et al., 2012; Lambert & Kormos, 2014; Larsen-Freeman, 2009; Norris & Ortega, 2009).

Complexity

Complexity is seen as the most controversial dimension of the three CAF constructs (Norris & Ortega, 2009; Pallotti, 2009, 2015). The confusion starts with the fact that complexity applies to different aspects of SLA. There is (1) developmental complexity (“the order in which linguistic structures emerge and are mastered in second (and, possibly, first) language acquisition” Pallotti, 2015, p. 2); (2) cognitive complexity (the subjective difficulty of a language feature, that is, how a learner perceives the difficulty of an item as it is processed and learned); and (3) linguistic complexity (objective complexity, which refers to “intrinsic formal or semantic-functional properties of L2 elements (e.g., forms, meanings and form–meaning mappings)” Housen et al., 2012, p. 4). To give an example, learners may perceive the English article system (zero, a/an, the) as very difficult, and its mastery might only follow at a later stage of development, while linguistically it could be argued to be fairly simple.

When measuring complexity, the linguistic dimension has most often been applied in CAF research. Linguistic complexity itself is a multidimensional construct. In their meticulous examination of L2 complexity, Bulté and Housen (2012, p. 24) define it as “the number of discrete components that a language feature or a language system consists of, and the number of connections between the different components.” They make a basic distinction between lexical complexity and grammatical complexity—a view that is in accordance with the body of empirical CAF studies.

Many scholars have set out to disentangle the different subdimensions of lexical complexity and to identify appropriate measures (e.g., Jarvis, 2013; Jarvis & Daller, 2013; Malvern & Richards, 1997; Vermeer, 2000). Most work differentiates lexical diversity (i.e., the size of the lexicon measured by means of, for example, type-token ratio measures), lexical sophistication (i.e., the depth of lexis measured by means of, for example, frequency of rare or academic words), and lexical density (i.e., the amount of information in a text, typically measured by the ratio of lexical words per function words). Bulté and Housen (2012) proposed to add compositionality (i.e., the number of formal and semantic components of lexical items) while Jarvis (2013) identified six subcomponents of lexical diversity: rarity, volume, variability, evenness, disparity, and dispersion. As can be imagined, providing an encompassing picture of the lexical complexity of L2 data is a challenging endeavor.

Key Concepts

Components of Lexical Complexity

- **Diversity:** Size of lexis; gauged by means of type-token ratio based measures.
- **Sophistication:** Depth of lexis; gauged by means of frequency measures, for example, of words beyond the 1,000 most common words.
- **Density:** Information packaging of lexis; gauged by means of, for example, ratio of lexical words per function words.

Components of Grammatical Complexity at Different Linguistic Levels (among Others Morphology, Syntax, Phonology)

- **Length:** Short versus long units; gauged by, for example, number of words per clause.
- **Variation:** Variety of units; gauged by, for example, number of different morphemes used.
- **Interdependence:** Relation between units; gauged by, for example, coordinated versus subordinated clauses.

Grammatical complexity, too, has different subdimensions. Even though most research has focused on syntactic complexity (sentence, clause, phrase), Bulté and Housen (2012) stress the importance of a morphological (inflectional, derivational) and phonological (suprasegmental, segmental) dimensions. At all these levels, one can distinguish less from more complex language in terms of length (e.g., longer sentences), variation (e.g., more frequent use of different types of morphemes), and interdependence (e.g., coordination versus subordination).

For both, lexical and grammatical complexity, the choice of which and how many components to employ and what exact measures to use is nontrivial as it will impact on the findings of empirical work. Norris and Ortega (2009) stress that one should avoid using co-linear measures (e.g., type-token ratio AND Guiraud's index because they both tap into lexical diversity). Instead, they suggest using measures that gauge different subcomponents and that are likely to distinguish between theoretically expected differences in the specific context. For example, to examine developmental changes at the syntactic level they propose measuring coordination (e.g., the number of coordinated phrases, a sign of complexification at initial stages of L2 proficiency), subordination (e.g., number of subordinate clauses, a good indicator of complexification at intermediate L2 levels), and phrase-internal complexification (e.g., length of noun phrases, for higher levels of L2 knowledge or L1 data). However, this development view has recently been challenged by Inoue (2016), who found task effects to be more important than proficiency. Similarly, Lambert and Kormos (2014) argue that these measures are not fine-grained enough and instead propose to analyze different types of subordination, for example, differentiating nominal subordination from subordination via subject/object relative clauses.

A final word of caution about complexity. There is a general tendency to interpret more complexity as an indicator of better language production, for example, more subordination should indicate higher levels of L2 use. However, as discussed by Pallotti (2009) this view is

Marije Michel

too simplistic. First, linguistic complexity varies by genre (e.g., small talk vs. argumentative essays) and individual stylistic preferences. To quote Pallotti (2009, p. 597): “Beckett is not Joyce, and this has nothing to do with (in)competence, but with stylistic choices.” Second, in a dynamic process like L2 development, linguistic complexity cannot be expected to grow linearly (Lambert & Kormos, 2014; Larsen-Freeman, 2009). As such, higher complexity (and also fluency) might indicate higher competence or performance levels, but this is by no means an absolute rule.

Accuracy

Accuracy seems to be the most transparent construct in the CAF triad (Housen & Kuiken, 2009; Pallotti, 2009; Wolfe-Quintero et al., 1998), and it refers to target-like-use of language, that is, error-free speech or writing, and measures the amount of deviation from the norm. The challenge of measuring accuracy is strongly related to the choice of linguistic norm, for example, a prescriptive grammatical description of the target language or native speaker usage. Applying a linguistic norm raises various issues, for example, a prescriptive norm might not be appropriate for spoken language use. Furthermore, raters do not always agree on what is accurate (cf. Kuiken & Vedder, 2014; Polio, 1997). The fact that the same language (e.g., German) may have several normative standards (e.g., Austrian, German, Swiss) adds another layer to this discussion.

Even if there was agreement regarding the norm, there remains the question of how ‘far away’ a deviation from this chosen norm is. For example, a punctuation error may not be as severe as mixing up word order, omitting an article, or using unusual lexical combinations as demonstrated by a comparison of (1) versus (2).

1. Honestly I think this is an excellent piece of writing.
2. Honestly, I think this tremendous writing is.

Valid and reliable measures of accuracy should be able to make this distinction (Polio & Shea, 2014). In this sense, Kuiken and Vedder (2008) distinguished first, second, and third degree errors in terms of communicative adequacy. Kuiken and Vedder’s (2008) categorization would classify (1) as a first degree minor error but (2) as more severe second degree error hampering understanding, while a third degree error would make the sentence incomprehensible. More recently, Foster and Wigglesworth (2016) have proposed a weighted measure for accuracy that assigns clauses a score based on their accuracy. Accordingly, the clauses in example (2) would receive the scores 1.0 for ‘Honestly, I think’ and 0.5 for ‘this tremendous writing is,’ for an overall score of 1.5. In this way, L2 production can be evaluated quantitatively by assigning a weighted single accuracy score to total performance. However, weighting errors reliably is not an easy task either (Pallotti, 2009), and as Foster and Wigglesworth (2016, p. 112) state: “Anyone who has worked on assessing accuracy in L2 data will know this only too well; some degree of personal judgment has to be invoked occasionally.”

In empirical work, accuracy has been gauged using holistic scales (e.g., Polio, 1997) and global measures (e.g., error-free clauses, number of errors per 100 words) as well as specific measures. The choice for a specific measure will be based on the language that is expected. For example, when investigating the effect of a teaching unit on past tense, target-like-use of past *-ed* would be the specific measure. Similarly, exploring language elicited by a task focusing on plural versus singular agents might count agreement errors, while the specific

L1-L2 combination could make it an obvious choice to go for gender marking on adjectives (for example, for English learners of Spanish).

Each measure comes with its own advantages and shortcomings. Holistic scales allow a global impression that takes into account the severity of errors; however, such scales often do not clearly distinguish accuracy from other dimensions such as complexity (Polio, 1997). Global measures make it possible to compare accuracy over different languages, populations, and tasks. Yet, they might not be sensitive enough to capture slight differences at higher levels of proficiency or of short-term interventions (Lambert & Kormos, 2014). In contrast, specific measures might be able to reveal small changes in accuracy, although it is difficult to generalize the findings to other contexts. Categorizing errors according to severity allows comparisons across studies, but it includes making strong interpretative choices when defining the categories and assigning an error to a certain degree.

Key Concepts

Measuring Accuracy

- **Holistic scales** provide a global impression of accuracy; for example, low score for “little knowledge of English vocabulary and word forms; virtually no mastery of sentence construction rules; dominated by errors” (Polio, 1997, p. 137).
- **Global** measures quantify overall accuracy; for example, number of error-free clauses.
- **Specific** measures focus on the specific target of a pedagogic intervention, task, or language; for example, number of noun-adjective-gender-agreement errors.
- **Degrees of errors** weight the severity of an error; for example, first degree: minor mistakes like spelling or omitted articles; second degree: more severe mistakes such as word order; third degree: mistakes that make an utterance nearly incomprehensible, e.g., combination of wrong word choice, word order and omissions (cf. Foster & Wigglesworth, 2016; Kuiken & Vedder, 2008).

To recap, even though accuracy seems to be less controversial than complexity, measuring this dimension of L2 use includes taking important decisions about the norm to choose and the severity of a deviance from this norm. In light of these considerations, Housen et al. (2012) appeal for using the abbreviation *A* not only for accuracy but also for appropriateness and acceptability, which would account for language use in different contexts and genres (e.g., *CU 2night* being appropriate in a text message but not in a formal invitation).

Fluency

In contrast to complexity and accuracy, which may pertain to oral and written L2 performance, fluency is first and foremost a measure of spoken language, even though writing research also uses measures of fluency. Historically and informally, the term fluency has been used to characterize a generally proficient L2 speaker (Chambers, 1997). More recent research adheres to a narrower definition (Lennon, 2000) where the construct is thought to encompass cognitive psychological, performative, and perceived aspects of fluency (Freed, 2000; Kormos & Dénes, 2004; Segalowitz, 2000, 2010). In ISLA, a definition by Tavakoli and Skehan (2005) is cited regularly, according to which fluency consists of the three subdimensions: (1) speed or rate, for example, number of words per minute; (2) silence or breakdown,

Marije Michel

that is, amount, location, and duration of (filled) pauses; and (3) repair, that is, false starts, repetitions, and self-corrections. In terms of language processing, speed is associated with control of and access to proceduralized knowledge; breakdown is thought to reflect the planning and conceptualization stages of language production; while repair fluency is seen as an indicator of monitoring processes (Levelt, 1989; Segalowitz, 2000, 2010; Skehan, 2003, 2009; Tavakoli & Skehan, 2005).

Key Concepts

Components of Fluency

- **Speed or rate:** Measured by, for example, syllables per second.
- **Silence or breakdown:** Measured by, for example, number, duration, and location (at clause boundaries vs. mid-clause) of pauses.
- **Repair:** Measured by, for example, number of false starts, repetitions, and self-repairs.

Measures of fluency based on temporal aspects of speech are relatively uncontroversial to identify and quantify in empirical research (Chambers, 1997), for example by calculating the ratio of syllables per second or the number of repairs per hundred words. It is important, however, to acknowledge that some aspects of fluency have been found to be trait-like personal characteristics rather than indicators of L2 competence (de Jong, Groenhout, Schoonen, & Hulstijn, 2015). De Jong, Steinel, Florijn, Schoonen, and Hulstijn (2012) advocate the use of phonation time ratio (“the percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample,” Kormos & Dénes, 2004, p. 148) instead of silence measures (see also Bosker et al., 2013 for a recent discussion). Moreover, Kormos and Dénes (2004) investigated the relationship between fluency measures and expert ratings of fluency, which revealed that boundaries between fluency, on the one hand, and complexity and accuracy on the other hand, are less clear-cut.

For writing, fluency is a more controversial construct because the reiterative process permits planning, monitoring, and editing (Johnson, Mercado, & Acevedo, 2012; Wolfe-Quintero et al., 1998). Typically, the oral measures of speed and breakdown are substituted by metrics of rate (e.g., number of words per minute based on the final text produced) and length (e.g., number of words per utterance). Yet, newer studies employed keystroke logging software (Leijten & van Waes, 2013; Révész, Kourтали, & Mazgutova, 2016) that records online writing features like number of characters typed between pauses or the ratio of number of characters produced during writing over the number of characters in the final text. Such measures make it less difficult to identify and disentangle the subdimensions of fluency from accuracy and complexity in written performance because they allow to review the process of writing fluency and not a product only.

To sum up, fluency is also a multifaceted construct with subcomponents. In particular in L2 writing, fluency constitutes a challenging dimension to measure and to conceptualize.

Measuring CAF

By now it has become clear that choosing measures of complexity, accuracy, and fluency needs careful considerations. Similarly, interpretations of results require caution and awareness of the explanatory power and limitations of the metrics used (Norris & Ortega, 2009).

In this chapter, no attempt is made to provide a list of the ‘best’ measures. Instead, some thoughts that guide the choice for or against a specific metric are shared.

Measures of CAF come in a variety of forms. Wolfe-Quintero et al. (1998) identify three types: (1) frequency counts of a specific linguistic unit, for example, number of word tokens; (2) ratio measures, that divide a specific unit by the total number of another unit, for example, type/token ratio (TTR); and (3) indices, which are calculations of a score by means of a more complex formula, for example, D is based on “mathematically modelling how new words are introduced into larger and larger language samples” (Malvern & Richards, 2002, p. 85).

The choice for a metric type will be based on the L2 data under investigation. For example, raw frequencies (e.g., total number of errors) allow comparisons only of L2 samples that are of equal length (e.g., texts of 300 words exactly). As soon as samples differ in length, ratios or indices should be used. Indices are calculated because some ratios are known to be nonlinearly affected by sample length (e.g., D by Malvern & Richards, 1997, 2002; or Measure of Textual Lexical Diversity, MTL D by McCarthy & Jarvis, 2010; both adjust TTR for sample length).

When calculating ratios and indices, an important decision is what unit of reference to use (e.g., sentences, clauses, words, minutes, seconds). While research into writing may count sentences (i.e., text between two period marks) as syntactic units, it is difficult to establish ‘sentence’ boundaries in oral performance. Alternative syntactic units include the terminal (T) unit (Hunt, 1965); the communication (C) unit (similar to T unit but including utterances without a verb; Bardovi-Harlig, 1992); and more recently the analysis of speech (AS) unit (Foster, Tonkyn, & Wigglesworth, 2000). The latter has become the standard for oral data (see also Crookes, 1990, for a discussion of different units).

Key Concepts

- **Terminal (T) unit:** (Hunt, 1965, p. 735): “One main clause plus whatever subordinate clauses happen to be attached or embedded within it.”
- **Analysis of Speech (AS) unit:** (Foster et al., 2000, p. 365): “A single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either.”

It is advisable to use to some extent the same measures as key references in earlier research to enable comparisons across studies. However, these should be supplemented by measures that are chosen specifically for the current study guided primarily by the type of data and the research questions. For example, Tonkyn (2012) employed eight specific structural measures because he examined development after a short-term intervention, and global measures may not have revealed a change. Michel (2013) decided to calculate the number of conjunctions per 100 words (and not per syntactic unit) in order to avoid interdependence of measures: conjunctions are used to introduce clauses and therefore correlate with the number of syntactic units. For practicality, de Jong et al. (2012) decided to exclude the location of pauses because they used an automatic script (de Jong & Wempe, 2009) to detect pauses in their analysis of over 2,000 speech samples and the script did not provide location information.

Marije Michel

To summarize, unless L2 samples are of exact equal length, it is advisable to employ ratios and indices rather than raw frequencies. Denominators for these ratios will differ across different measures. Grammatical complexity and accuracy are traditionally expressed as a ratio per syntactic unit (e.g., errors per AS unit). Lexical measures typically take as denominator the number of words (tokens), while fluency employs temporal units such as minutes.

Empirical Evidence

CAF measures have been used to examine L2 performance, proficiency, and development in a wide variety of fields, including work investigating learner internal factors such as personality (e.g., Dewaele & Furnham, 2000) and age (e.g., Munoz, 2006), as well as external factors such as a specific instructional interventions (e.g., Derwing & Rossiter, 2003; Tavakoli, Campbell, & McCormack, 2015), the learning context (Housen et al., 2011; Mora & Valls-Ferrer, 2012), and many others. This section presents a selective review of empirical work with a focus on studies into different task design and task condition factors that can be manipulated in the classroom. Finally, some work that has used a longitudinal design is presented to provide a developmental perspective.

Task Complexity

Task complexity, that is, the cognitive demands of a task, has received ample attention over the past two decades, particularly in empirical research investigating the claims of Robinson's (2001) Cognition Hypothesis and Skehan's (1998) Limited Attentional Capacity Model. In short, Skehan predicts that higher cognitive task demands will inevitably result in trade-off effects, in particular between complexity and accuracy, due to competition for limited attentional resources (see Skehan, 2009, for the rationale based on Levelt, 1989). On the contrary, Robinson claims that parallel increases of complexity and accuracy are possible under certain conditions of task design (e.g., when a task requires more reasoning) because the higher cognitive demands require more focused linguistic performance.

Over the years, many studies have set out to contribute to the debate (e.g., the studies gathered in Robinson, 2011). Yet, as Jackson and Suethanapornkul's (2013) research synthesis shows, no compelling answers have been found due, in part, to the large variety of research designs and a plethora of CAF measures generating conflicting results. A meta-analysis of nine comparable studies (Jackson & Suethanapornkul, 2013) revealed that an increase of task complexity resulted in small positive effects for accuracy and small negative effects for fluency (a finding that is consistent with both hypotheses, cf. Skehan, 2009) while grammatical complexity was affected negatively and lexis positively. The latter two findings, however, were not robust enough to support or reject either of the two claims. By synthesizing the findings of seven of their earlier investigations, Skehan and Foster (2012) come to a similar conclusion, that is, they cannot present firm generalizations because the variety of instruments and measures offered different information.

It is in light of these inconclusive findings from numerous studies that Long (2016) reiterates Norris and Ortega's (2009) call for more standardization and a unified approach to the investigation of task complexity in the future.

Teaching Tips

- It is important to be aware that L2 users are likely to be less fluent when confronted with more complex tasks. However, the higher cognitive demands are likely to result in more accurate and/or complex language, and instructors and learners can monitor these to evaluate progress.
- Task repetition and familiarity is a fruitful way to foster higher levels of performance in terms of CAF. Repeating a task just once may enhance students' fluency. If targeting accuracy and complexity, multiple task repetitions might be needed to let students overcome trade-off effects between these two dimensions.
- Planning time can be given before (strategic pretask) or during (unpressured within-task) performance. Giving pretask planning time is likely to increase complexity and fluency because L2 speakers can conceptualize their performance beforehand. Giving students time to perform a task at their own pace (within-task planning time) decreases fluency but will positively affect complexity and/or accuracy (presumably not both due to trade-off effects).

Task Repetition and Familiarity

More systematicity in experimental design might be found in the body of research (e.g., Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; Kim & Tracy-Ventura, 2013; Mackey, Kanganas, & Oliver, 2007; Pinter, 2005) that looked into effects of task familiarity and task repetition, that is, “repetitions of the same or slightly altered tasks—whether whole tasks, or parts of a task” (Bygate & Samuda, 2005, p. 43). Many of these investigations employed CAF measures to evaluate L2 performance. Accordingly, when adults and young learners performed the same or a familiar task more than once, they were more fluent. Findings for complexity and accuracy have resulted in less clear patterns. As Bygate and Samuda (2005) hypothesize, repeated encounters allow L2 performers to shift from meaning-oriented toward more form-oriented production, the latter potentially creating trade-off effects between linguistic complexity and accuracy (Skehan, 2009). Overall, however, students' performance seems to improve when they work more than once on the same or similar material and CAF scores increase accordingly. Using slightly different content for similar tasks (i.e., task familiarity) seems to sustain students' motivation and interest over multiple repetitions.

Planning Time Studies

Also, providing L2 users with planning time seems to lead to higher levels of performance, in particular with respect to fluency. Effects of planning time (pretask planning, within-task planning, task rehearsal) on CAF has been extensively investigated and includes work into oral as well as written production (e.g., Ellis & Yuan, 2005; Foster & Skehan, 1999; Ortega, 2005). Mehnert (1998) showed that an absence of planning time resulted in low fluency scores, while different lengths of pretask planning time seemed not to make a difference. In his introduction to an edited volume on planning, Ellis (2005) summarizes that strategic pretask planning positively affects complexity and fluency while effects on accuracy are mixed. Skehan and Foster's (2012) synthesis of their earlier work indicates that pretask

Marije Michel

planning time affects the conceptualizing stage of speech performance and, therefore, promotes mainly structural complexity and lexical sophistication but also accuracy. The various aspects of fluency (speed, pauses, repair) were found to be affected in different ways.

Recently, Hsu (2015) looked into the effects of planning time in written synchronous computer-mediated communication (SCMC or text chat). Pretask rehearsal planning time was operationalized as writing a picture description during 10 minutes immediately before ‘telling’ that story to an SCMC interlocutor. Results showed that rehearsal planning time increased accuracy while complexity seemed to be unaffected. Regarding unpressured, within-task planning, the studies gathered in Ellis (2005) indicate that it promotes accuracy and also complexity while fluency decreases, a finding that was recently replicated by Ahmadian (2012).

To summarize, planning time seems to support conceptualizing (pretask) and monitoring (within-task), which has the potential to lead to higher scores on all three CAF dimensions. However, trade-off effects are likely to become visible, for example, increased accuracy as a result of monitoring during within-task planning time might come at the cost of fluency.

Modality: CAF in Oral Versus Written Versus Computer-Mediated Communication

In contrast to the large amount of work on planning time, there is only a handful of CAF studies that has explored effects of different modalities (oral, written, computer-mediated) on L2 performance. Using a between-participant design, Kuiken and Vedder (2012) compared oral versus written production at different levels of task complexity. Their results showed only minor differences between the two modalities. Ellis and Yuan (2005) looked at effects of planning conditions in oral versus written performances. They found greater complexity and accuracy but lower fluency in writing, which they attributed to the fact that writing allows for more planning, formulating, executing, and monitoring than speaking.

Sauro (2012) compared oral and written SCMC interactions of L2 speakers. Using measures of complexity and accuracy, no significant differences between the two modes could be attested in group comparisons. Yet, in individual evaluations, large variation between participants emerged, which Sauro assigned to discourse style and turn-taking behavior as well as typing skills.

In sum, these studies seem to suggest that CAF is not so much affected by modality apart from the expected effects of increased planning time and monitoring during writing.

Teaching Tips

- Be aware that ‘more’ (e.g., complex, fluent) does not automatically entail ‘better.’
- In addition to CAF, there are good reasons to measure performance in terms of communicative adequacy and task completion.
- Some aspects of language use have shown to be related to individual characteristics of a speaker (e.g., syllable duration) and/or are elicited by a specific genre or task feature (clause length). Therefore, such features may not be suitable indicators of proficiency.
- Lack of improvement on one dimension does not mean there is no improvement. Many studies suggest trade-off effects between complexity, accuracy, and fluency (de Jong et al., 2015; Pallotti, 2009).

Longitudinal Development

The majority of developmental CAF research has looked into the three dimensions using cross-sectional designs, with just a few recent longitudinal studies.

In writing research, Spoelman and Verspoor (2010) used analytical tools from dynamic systems theory (DST; e.g., Monte Carlo simulations) to explore 54 writing samples of a single learner studying Finnish during 3 years. Although complexity and accuracy of Finnish case marking showed growth over time, development was nonlinear. That is, the data revealed peaks, regressions, and backsliding on specific dimensions, as well as complex interactional patterns among the three dimensions. In another study, Gunnarsson (2012) followed the development of CAF in the written performance of five Swedish L2 learners of French over a period of 30 months. Analyses revealed large individual differences pointing to trade-off effects (Skehan, 2009), that is, while some writers showed gains in accuracy at the expense of fluency, others prioritized fluency at the cost of accuracy. Polio and Shea (2014) focused on the development of accuracy in a corpus of ESL learners who received writing instruction over the course of one semester (Polio, 1997). They found minor improvements of accuracy but increased linguistic complexity, which they interpret as a trade-off effect. The corpus-based study by Vyatkins, Hirschmann, and Golcher (2015) used multilevel modeling to investigate syntactic development of seven different modifiers (e.g., adverbs, prepositional phrases) in longitudinal writing data of English learners of German over the course of four semesters. This study showed that the global use of modifiers remained relatively stable but the type of modification revealed large inter- and intra-individual variation over time.

Investigations into oral performance include Ferrari (2012), who looked into the development of CAF in four adolescent L2 learners and two native speakers of Italian who performed monologic and dialogic tasks over the course of 3 years. Her findings suggested trade-off effects between different CAF components in different communicative situations but generally, monologic tasks created greater complexity but lower fluency than dialogic performances. Based on a detailed comparison of the L2 and L1 data, Ferrari concluded that “the ability to vary one’s language according to the demands of different communicative activities develops very slowly” (p. 294). In contrast, Vercellotti (2015) could not detect trade-off effects in her data on the oral performance of 66 L2 learners who were recorded monthly over a period of 10 months during an intensive English program. Using hierarchical linear modeling she found that grammatical complexity, accuracy, and fluency showed steady linear growth while lexical variety revealed a nonlinear trajectory, that is, there was a dip followed by a steep increase. Finally, in the case study by Polat and Kim (2014) one uninstructed L2 speaker was interviewed biweekly during a full year and they used dynamic systems theory methods to gain insights into complexity and accuracy development. While lexical complexity showed steady growth over time and syntactic complexity somewhat increased, accuracy seemed unaffected.

Future Directions

The Role of Communicative Adequacy

Even though research into CAF suggests that the triad appropriately captures relevant aspects of L2 performance, a call for the inclusion of communicative or functional adequacy has been issued more than once in recent years. Pallotti (2009, p. 596) defines this fourth construct as “the degree to which a learners’ performance is more or less successful in

achieving the task's goals efficiently." For instance, an utterance scoring high on all three CAF measures can be communicatively inadequate and vice versa, which shows the independence of the two constructs. In language pedagogy and testing, communicative adequacy is one of the main goals, as evidenced for example by the Common European Framework of Reference (CEFR).

To date, only a handful of studies have looked at CAF and communicative adequacy, revealing that they are complementary constructs interacting in several ways. Kuiken, Vedder, and Gilabert (2010) showed that adequacy ratings on L2 writing were not so much correlated to structural complexity, while lexical complexity and accuracy were. Révész, Ekiert, and Torgersen (2014) employed linear mixed effects regression and Rasch analyses to investigate adequacy in spoken performance. In their data, the number of filled pauses (i.e., breakdown fluency) seemed to be the strongest predictor of communicative adequacy, while other CAF measures showed minor effects only. Yet another study (de Jong et al., 2012) identified vocabulary knowledge and correct sentence intonation as the strongest predictors of adequacy by means of structural equation modeling.

CAF in Interaction

A disregarded issue in past research has been how the CAF triad accounts for differences between dialogic and monologic performance (but see Ferrari, 2012, reviewed earlier). Among the few studies, Michel, Kuiken, and Vedder (2007) and Michel (2011) gave the same tasks to L2 (and L1) speakers of Dutch working either on their own or in pairs. Dialogic performance in both populations was characterized by lower grammatical complexity, but higher accuracy and fluency. While nonnative speakers were lexically more varied, native speakers showed lower lexical variety in dialogues. Similarly, Gilabert, Barón, and Levkina (2011) found dialogic performances to be more fluent but grammatically less complex.

From a methodological perspective, these studies raise the question of whether current CAF measures gauge the same constructs in monologues and dialogues and whether measurement is valid and reliable. Indeed, both Sato (2014) and Tavakoli (2016), who focus on fluency, state that we might need other measures in dialogues that account for interactive turn-taking patterns because fluency in individual versus interactional performance is fundamentally different. Tavakoli compared several established and newly developed measures of fluency when evaluating monologic and dialogic L2 speech. Findings showed that well-known fluency metrics for monologic production (e.g., phonation time ratio) might not be reliable measures in dialogue, because overlapping speech and between-speaker pauses need to be divided over partners. Earlier, Sato (2014) had already established that raters take into account effective scaffolding and disruptive pause behavior in dialogic speech when assigning fluency scores to speech samples.

Measuring Instructional Effects by Means of CAF

To date surprisingly few studies have used CAF to gauge instructional effects. One reason could be the earlier mentioned concern that global CAF measures might not be sensitive enough to capture slight differences of performance after (short-term) pedagogical interventions. Another cause could be the fact that many interventions focus on a specific linguistic target and, therefore, structure-focused pretest/posttests—rather than global CAF performance measures—are thought to be more suitable. Exceptions are the aforementioned work by Mora and Valls-Ferrer (2012), Tavakoli et al. (2015), as well as Tonkyn (2012)

and other chapters in the edited volumes by Housen et al. (2012) and Baralt, Gilabert and Robinson (2014). With the development of more fine-grained measures (for example, the ones proposed by Lambert & Kormos, 2014, for syntactic and by Jarvis, 2013, for lexical complexity, respectively) and scores (for example, Foster & Wigglesworth, 2016, weighted accuracy score) future work will hopefully aim to capture instructional effects by means of CAF. The use of CAF measures in future ISLA studies might be further promoted by the growing availability of computerized tools that provide fast and reliable ways to measure CAF. The next section will highlight a few of these tools, knowing the risk of obsolescence due to rapid developments in this area.

Computer-Based Tools and Corpus-Based Techniques for Analyzing CAF

For syntactic complexity, Coh-Metrix (McNamara, Louwerse, Cai, & Graesser, 2013) and Synlex (Lu, 2010), which produce output metrics for length of syntactic units as well as coordination, subordination and syntactic sophistication in L2 writing are widely used. Many (web-based) tools exist that provide calculations of type/token ratios and other measures of lexical diversity, sophistication and density (among others AntWordProfiler, Anthony, 2015; LexTutor for English and French, Cobb, 2002). Fortunately, language corpora are often error-tagged, which allows automatic accuracy measurement. However, automatic computer-based accuracy measurement remains a desideratum.

Teaching Tip

- Let students use software (e.g., Synlex and LexProfiler) to analyze the changing complexity of their writing, for example, over tasks, genre and time. Exploring complexity is likely to raise their awareness that accuracy is only one aspect of L2 performance. That is, it might help them to realize that they are making progress in terms of complexity even though error rates do not suggest development.

For fluency, CLAN (MacWhinney, 2000) and Praat (Boersma & Weenink, 2013) are widely used. The language independent Praat-script that counts the number of syllables and silent pauses (de Jong & Wempe, 2009) is particularly relevant for the fluency analysis of oral data. Keystroke logging programs such as InputLog (Leijten & van Waes, 2013; see also <http://www.writingpro.eu>) allow the investigation of speed, pause and revision measures in computer-based writing, which promises to boost future work into L2 writing by means of CAF.

Finally, corpus-based research facilitates the analysis of developmental trajectories based on large amounts of data (Alexopoulou, Michel, Murakami, & Meurers, 2017; Thewissen, 2013; Vyatkina et al., 2015). As more corpora and tools for different languages become available, computer-based CAF research faces a promising future. In the same vein, it is hoped that future longitudinal research will be able to uncover further developmental patterns and individual trajectories of CAF in oral and written L2 performance. From a pedagogic perspective, more future work is needed into the complex interrelationship between communicative adequacy and CAF, in particular, in dialogic settings given that L2 instruction often involves pair work.

Marije Michel

Conclusion

Researchers seem to agree that the CAF triad is a useful and valid way to investigate and describe L2 performance and development. However, to date, no consensus has been reached on how to define and measure the constructs.

Over the past decades, many have set out to identify the ‘best’ or a ‘better’ measure (e.g., Kormos & Dénes, 2004; Pallotti, 2015; Polio, 1997; Wolfe-Quintero et al., 1998). Although these investigations add to our knowledge and understanding, a result is that there are a daunting number of metrics available. For example, Long (2015) criticizes the fact that 84 different measures have been used to examine effects of task complexity. In addition, little is known about the validity and reliability of many measures because most research has paid little attention to these issues. As outcomes are based on different metrics of unknown reliability and validity, it is difficult to identify general trends and compare findings. Consequently, the future calls, on the one hand, for greater standardization and theory-driven use of constructs and metrics and, on the other hand, for the acknowledgment of variability and dynamicity of CAF in L2 language use (Housen et al., 2012; Norris & Ortega, 2009).

Acknowledgments

I would like to thank Gabriele Pallotti and the editors who all gave valuable feedback on previous versions of this chapter.

References

- Ahmadian, M.J. (2012). The effects of guided careful online planning on complexity, accuracy and fluency in intermediate EFL learners’ oral production: The case of English articles. *Language Teaching Research*, 16(1), 129–149.
- Ahmadian, M.J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners’ oral production. *Language Teaching Research*, 15(1), 35–59.
- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D. (2017, Advance Access) Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing Natural Language Processing techniques. *Language Learning*, Special Issue on Language learning research at the intersection of experimental, corpus-based and computational methods.
- Anthony, L. (2015). AntWordProfiler (Version 1.4.1) [Computer program]. Tokyo, Waseda University. Retrieved from <http://www.laurenceanthony.net/>
- Baralt, M., Gilabert, R., & Robinson, P. (Eds.). (2014). *Task sequencing and instructed second language learning*. London: Bloomsbury.
- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26(2), 390–395.
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Retrieved from <http://www.praat.org/>
- Bosker, H., Pinget, A., Quené, H., Sanders, T., & de Jong, N. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, 30, 159–175.
- Brumfit, C. (1979). Communicative language teaching: An educational perspective. In C.J. Brumfit & K. Johnson (Eds.), *The communicative approach to language teaching* (pp. 183–191). Oxford: Oxford University Press.
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 23–46). Amsterdam/Philadelphia: John Benjamins.

- Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). Oxford: Macmillan Heinemann.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In K. van den Branden, M. Bygate, & J. Norris (Eds.), *Task-based language teaching: A reader* (Vol. 1, pp. 249–274). Amsterdam/Philadelphia: John Benjamins.
- Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task-repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37–74). Amsterdam/Philadelphia: John Benjamins.
- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535–544.
- Cobb, T.: Web Vocabprofile, an adaptation of Heatley, Nation & Coxhead's (2002). Range. [Computer program]. Retrieved from <http://www.lex tutor.ca/vp/>
- Crookes, D. (1990). The utterance and other basic units for second language discourse analysis. *Applied Linguistics*, 11, 183–199.
- de Jong, N.H., Groenhout, R., Schoonen, R., & Hulstijn, J.H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 36, 223–243.
- de Jong, N.H., Steinel, M., Florijn, A., Schoonen, R., & Hulstijn, J. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5–34.
- de Jong, N.H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, 41(2), 385–390.
- Derwing, T.M., & Rossiter, M.J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, 13(1), 1–17.
- Dewaele, J.M., & Furnham, A. (2000). Personality and speech production: A pilot study of second language learners. *Personality and Individual Differences*, 28(2), 355–365.
- Ellis, R. (Ed.). (2005). *Planning and task performance in a second language*. Amsterdam/Philadelphia: John Benjamins.
- Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task-based performance in a second language* (pp. 167–192). Amsterdam/Philadelphia: John Benjamins.
- Ferrari, S. (2012). A longitudinal study of complexity, accuracy and fluency variation in second language development. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp. 277–298). Amsterdam/Philadelphia: John Benjamins.
- Foster, P., & Skehan, P. (1999). The effect of source of planning and focus on planning on task-based performance. *Language Teaching Research*, 3(3), 185–215.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–375.
- Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36, 98–116.
- Freed, B. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder. In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 243–265). Ann Arbor, MI: University of Michigan Press.
- Gilbert, R., Barón, J., & Levkina, M. (2011). Manipulating task complexity across task types and modes. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 105–138). Amsterdam/Philadelphia: John Benjamins.
- Gunnarsson, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp. 247–276). Amsterdam/Philadelphia: John Benjamins.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461–473.
- Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy, and fluency in SLA*. Amsterdam/Philadelphia: John Benjamins.

Marije Michel

- Housen, A., Schoonjans, E., Janssens, S., Welcomme, A., Schoonheere, E., & Pierrard, M. (2011). Conceptualizing and measuring the impact of contextual factors in instructed SLA—the role of language prominence. *IRAL—International Review of Applied Linguistics*, 49(2), 83–112.
- Hsu, H. C. (2015, advance access). The effect of task planning on L2 performance and L2 development in text-based synchronous computer-mediated communication. *Applied Linguistics*, 1–28.
- Hunt, K. W. (1965). *Grammatical structure written at three grade levels* (Research Report 3). Urbana, IL: National Council of Teachers of English.
- Inoue, C. (2016). A comparative study of the variables used to measure syntactic complexity and accuracy in task-based research. *Language Learning Journal*, 44(4), 1–19.
- Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, 63(2), 330–367.
- Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning*, 63(s1), 87–106.
- Jarvis, S., & Daller, M. (Eds.). (2013). *Vocabulary knowledge: Human ratings and automated measures*. Amsterdam/Philadelphia: John Benjamins.
- Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, 21, 264–282.
- Kim, Y. J., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, 41(3), 829–840.
- Koponen, M., & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5–24). Ann Arbor: University of Michigan Press.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164.
- Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, 17(1), 48–60.
- Kuiken, V., & Vedder, I. (2012). Syntactic complexity, lexical variation and accuracy as a function of task complexity and proficiency level in L2 writing and speaking. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 143–169). Amsterdam: John Benjamins.
- Kuiken, F., & Vedder, I. (2014). Rating written performance: What do raters do and why? *Language Testing*, 31(3), 329–348.
- Kuiken, F., Vedder, I., & Gilabert, R. (2010). Communicative adequacy and linguistic complexity in L2 writing. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 81–100). Eurosla Monographs 1. Eurosla.
- Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics*, 35(5), 607–614.
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly*, 12(4), 439–448.
- Larsen-Freeman, D. (2009). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 30(4), 590–619.
- Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using InputLog to analyze and visualize writing processes. *Written Communication*, 30(3), 358–392.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387–417.
- Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25–42). Ann Arbor, MI: University of Michigan Press.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Oxford: Wiley-Blackwell.
- Long, M. H. (2016). In defense of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics*, 36, 5–33.

- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474–496.
- Mackey, A., Kanganas, A., & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly*, 41(2), 285–312.
- MacWhinney, B. (2000). The chldes project: Tools for analyzing talk [Computer Program]. Retrieved from <http://chldes.psy.cmu.edu/>
- Malvern, D., & Richards, B. (1997). A new measure of lexical diversity. In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 58–71). Clevedon: Multilingual Matters.
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85–104.
- McCarthy, P. M., & Jarvis, S. (2010). MTL-D, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392.
- McNamara, D. S., Louwerse, M. M., Cai, Z., & Graesser, A. (2013). Coh-Metrix version 3.0. Retrieved from <http://cohmetrix.com>
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20(1), 52–83.
- Michel, M. (2011). Effects of task complexity and interaction on L2 performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 141–174). Amsterdam/Philadelphia: John Benjamins.
- Michel, M. (2013). Effects of task complexity on the use of conjunctions in oral L2 task performance. *Modern Language Journal*, 97(1), 178–195.
- Michel, M., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, 45(3), 241–259.
- Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46(4), 610–641.
- Munoz, C. (Ed.). (2006). *Age and the rate of foreign language learning*. Clevedon: Multilingual Matters.
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578.
- Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77–109). Amsterdam/Philadelphia: John Benjamins.
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590–601.
- Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, 31(1), 117–134.
- Pinter, A. (2005). Task repetition with 10-year-old children. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 113–126). Basingstoke: Palgrave Macmillan.
- Polat, B., & Kim, Y. J. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, 35(2), 184–207.
- Polio, C. G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, 47, 101–143.
- Polio, C. G., & Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10–27.
- Révész, A., Ekiert, M., & Torgersen, E. (2014, advance access). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics*, 1–22.
- Révész, A., Kourтали, N., & Mazgutova, D. (2016, advance access). The effects of task complexity on L2 writing processes, behaviours and linguistic complexity. *Language Learning*.
- Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27–57.
- Robinson, P. (Ed.). (2011). *Second language task complexity: Researching the cognition hypothesis of language learning and performance*. Amsterdam/Philadelphia: John Benjamins.

Marije Michel

- Sato, M. (2014). Exploring the construct of interactional oral fluency: Second language acquisition and language testing approaches. *System*, 45, 79–91.
- Sauro, S. (2012). L2 performance in text-chat and spoken discourse. *System*, 40, 335–348.
- Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor: University of Michigan Press.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. London: Routledge.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(1), 1–14.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532.
- Skehan, P., & Foster, P. (2012). Complexity, accuracy, and fluency and lexis in task-based performance: A synthesis of the Ealing research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp. 199–220). Amsterdam/Philadelphia: John Benjamins.
- Spelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31(4), 532–553.
- Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, 54(2), 133–150.
- Tavakoli, P., Campbell, C., & McCormack, J. (2015). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Quarterly*, 50(2), 1–25.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–277). Amsterdam/ Philadelphia: John Benjamins.
- Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *Modern Language Journal*, 97, 77–101.
- Tonkyn, A. (2012). Measuring and perceiving changes in oral complexity, accuracy and fluency. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency* (pp. 221–245). Amsterdam/Philadelphia: John Benjamins.
- Vercellotti, M. L. (2015, advance online access). The Development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics*, 1–23.
- Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17(1), 65–83.
- Vyatkina, N., Hirschmann, H., & Golcher, F. (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing*, 29, 28–50.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (No. 17). Honolulu: University of Hawai'i Press.

Sociocultural Theory in the L2 Classroom

Neomy Storch

Background

Historical Background

Sociocultural theory of mind, commonly abbreviated to SCT, is based on the work of Soviet psychologist Lev Vygotsky (1978, 1981). It was further developed by his Soviet colleagues (e.g., Leontiev, 1978), as well as by Western scholars in the field of psychology and education (e.g., Wells, 1999; Wertsch, 1991). In applied linguistics, the first to employ SCT was James Lantolf. In a series of studies he conducted with Frawley (e.g., Frawley & Lantolf, 1985), SCT was employed to examine how second language (L2) speakers use their L2 to mediate their performance when completing difficult tasks.

SCT views the development of all complex human cognitive facilities, including the learning of first and subsequent languages (Luria, 1973), as inherently social and mediated by artefacts (e.g., texts, gestures). Initially SCT was met with vigorous resistance from established researchers in the field of SLA (e.g., Gregg, 1993; Long, 1990), who opposed the proliferation of theories attempting to explain SLA and particularly theories, such as SCT, which view language learning as a social rather than a purely cognitive phenomenon. However, over the years, the theory has become a more accepted perspective in mainstream SLA and L2 pedagogy. The growing acceptance of SCT in applied linguistics research is evident in numerous chapters in edited books and handbooks on SLA, articles in leading journals in the field, and PhD dissertations all using SCT to address issues pertinent to L2 learning, teaching, and testing. Much of this acceptance is no doubt due to the work of scholars such as Lantolf (e.g., Lantolf, 2000; Lantolf & Thorne, 2006), who has explicated key theoretical constructs of relevance to L2 learning and teaching, and to Swain (e.g., Swain, 2000, 2006; Swain, Kinneer, & Steinman, 2011), who has made these constructs much more concrete and accessible to SLA researchers by showing how they can be used to explain the nature and focus of L2 learners' interaction. Swain's research has also shown the inseparability of the social and cognitive dimensions of L2 learning.

Two strands in this growing body of research can be discerned. One strand uses SCT to provide a rationale for and to explain findings from studies investigating various classroom activities (e.g., pair work) and teacher interventions (e.g., feedback). Although the use of SCT for this purpose has been criticised by some scholars (e.g., van Compernelle & Williams, 2013), there is a growing body of research (see subsequent discussion) that has employed SCT for this purpose. The other strand uses SCT as a theoretical framework in the design of a coherent teaching program (e.g., concept-based instruction) or in the design of assessment practices (e.g., dynamic assessment). The aim of both strands of research is to inform L2 pedagogy.

The studies discussed in this chapter reside in the first strand of this research. The chapter begins with a brief overview of SCT. Although SCT has a number of central constructs, the constructs that are the focus of this chapter are the zone of proximal development (ZPD) and mediation due to their relevance to L2 instruction. They provide a rationale for certain approaches to teacher interventions (e.g., feedback) and for certain classroom activities (e.g., pair/group work), a rationale that differs to that provided by psycholinguistic and cognitive theories of SLA.

Overview of SCT

It is important to acknowledge at the outset that SCT is not a theory of second language learning but rather a psychological theory that explains how biologically endowed mental capacities (e.g., memory, involuntary attention) develop into uniquely human higher order cognitive capacities (e.g., intentional memory, voluntary attention, planning), over which humans, unlike other species, can exercise control. The underlying premise in SCT is that the development of these higher order cognitive capacities occurs in contextualised interactions between an expert member of the community (e.g., an adult, a knowledgeable peer) and a novice (e.g., a child, a less knowledgeable peer). These interactions are mediated by tools, which may be physical (e.g., computers) or symbolic (e.g., gestures, language). These tools enable interaction to take place (e.g., via dialogue, use of gestures, or use of computer-mediated forms of communication); they also enable humans to solve problems and to develop higher order capacities.

Unlike other psychological theories of cognitive development (e.g., Piaget, 1977), SCT views the direction of cognitive development from the social to the individual. SCT proposes that cognitive functions appear first in social interactions between humans, and that they subsequently become internalised within the individual. This development is perceived as increasing regulation: the novice transforms from being object-regulated (reliant on concrete physical representation of objects such as the reliance on realia in beginner L2 classes), to being other-regulated (reliant on the assistance of an expert such as the teacher or textbooks to produce and comprehend the L2) to ultimately becoming self-regulated (independent user of the L2 who is able to rely on abstract rules of the L2 when producing and comprehending language). However, it should also be noted that SCT does not view internalisation as a process whereby the novice simply imitates the expert (Lantolf & Thorne, 2006), but rather as a transformative process. The novice processes the knowledge that was co-constructed with the expert and makes it her own unique resource. Knowledge in this sense (including language knowledge) is not an object to be possessed or accumulated by the individual (see Sfard, 1998), but an understanding that is “recreated, modified, and extended in and through collaborative knowledge building and individual understanding” (Wells, 1999, p. 89).

The Zone of Proximal Development (ZPD)

From an SCT perspective, development occurs in interaction between an expert and a novice, where the expert provides assistance to the novice. However, not all assistance provided by the expert is supportive of development. As Lantolf and Thorne (2006) point out, some forms of assistance may be inappropriate and constrain development. Too much assistance (as well as too little assistance) may be detrimental to development. Effective assistance needs to be contingently responsive to the learner's need for assistance, and ultimately involve a "handover" of responsibility to the learner (van Lier, 2004), so that the learner can perform tasks independently.

Appropriate levels of assistance not only promote development but can also be used to measure development. The type of assistance that a novice needs to complete a task is, according to Vygotsky (1978), more indicative of the novice's potential development than unassisted, independent performance. Vygotsky explains that independent performance merely measures the novice's current capacity; performance with assistance measures capacity for cognitive growth. When comparing the performance of two novices, the novice who can take advantage of assistance is judged as having a greater potential for cognitive development than one who cannot take advantage of the assistance offered. This notion of potential development as distinct from the current level of performance is encapsulated in the ZPD.

In formal instructed settings, ZPD implies that effective instruction should be forward looking (Vygotsky, 1978) and attuned to the learner's potential maturing capacities rather than being based on fully matured capacities evident in current performance. Thus it implies that the expert needs to monitor the learner's ability to take advantage of the assistance provided and of any changes in this ability. In other words, effective assistance is dynamic. Another important trait of the ZPD is that it is co-constructed by the expert and the learner (Roth & Radford, 2010), and that the contributions of the learner are very important because they provide a cue to the expert. Poehner (2008), for example, writing about the reciprocity of the learner in the ZPD, notes that we should view the learner as agentive rather than as a passive recipient of assistance. The support offered needs to make challenging tasks accessible but also encourage learner engagement.

The metaphor that has been used in the literature to describe this contingently responsive and dynamic assistance is scaffolding, a term first introduced by Wood, Bruner, and Ross (1976) to describe child-adult (tutor) interaction. The appeal of this metaphor, as Wilson and Devereux (2014) suggest, is that it conjures up the idea of learning as a building under construction. The scaffold is vital for the construction to take place, but it is a temporary structure. As the construction progresses, the scaffold is gradually dismantled and it is removed when the building can stand alone. In education, scaffolding should also be perceived as an important temporary structure. It enables the learner to perform a task beyond their current capacity, but it should be gradually dismantled in line with the learner's increasing expertise and removed when it is no longer needed; that is, when the learner can complete the task independently.

Mediation

Mediation is another key construct in SCT. Mediation occurs when we use tools to enable or enhance our actions, including our thinking processes (Vygotsky, 1978). These tools can be physical (material) artefacts such as textbooks and computers or symbolic such as signs, gestures, and language. Of all symbolic tools, language is considered the most powerful

mediating tool but only when it is used for cognitive purposes (e.g., to plan, to focus attention) rather than for social purposes (e.g., to greet people) (Vygotsky, 1978; Vygotsky & Luria, 1994).

As a cognitive tool, language enables actions to take place between and within individuals. Between individuals, language is other-directed, social speech. It enables the novice and expert to communicate and coordinate their action (Wells, 1999), to invoke and share attention, and to co-construct the scaffold. Within individuals, language enables the individual to structure and organise their actions, including their thinking processes. This form of self-directed speech is termed “private speech.” It emerges when we engage in complex tasks, is often subvocal (whispered), and takes the form of short, incomplete phrases (Swain et al., 2011). Private speech enables us to focus our attention, retrieve stored information, and assess this information. By transforming our thoughts into words, our thoughts become artefacts that can be further reflected upon. At the same time, articulating our thoughts, whether to ourselves or to others, helps us to gain a deeper understanding of complex phenomena and to solve problems. In this sense other- and self-directed speech are important processes and products.

Swain (2006) proposed the term “*linguaging*” to describe how language mediates the thinking process. She defined *linguaging* as a “process of making meaning and shaping knowledge and experience through language” (2006, p. 98). *Linguaging* can be via speaking or writing. For L2 learners, *linguaging* can occur in the L1, the target language, or indeed any other additional language that the learner has learnt. *Linguaging* can be within the individual (self-directed private speech) as well as between individuals (collaborative talk).

It should be noted that the reference to collaborative talk in SCT is quite distinct from the notion of negotiation of meaning in Long’s (1985) interaction hypothesis. Negotiation of meaning occurs when learners experience or anticipate some sort of communicative breakdown because the input is incomprehensible. Negotiation of meaning between the interlocutors aims to make the input comprehensible, which is said to lead to L2 learning. In contrast, the purpose of collaborative talk is not to make input more comprehensible, but rather to solve a problem. As learners deliberate about how to solve a problem at hand, they draw on their own as well as each other’s linguistic resources; they build and extend on these resources in what has been labelled “collective scaffolding” (Donato, 1994; Storch, 2002). In the process, new knowledge is created or existing knowledge is consolidated and extended. Collaborative talk enables learners to reach resolutions to language-related problems that they may not have been able to reach had they been working on their own.

Current Issues

Research informed by SCT can address a number of current issues related to L2 teaching and learning. One such issue is the nature of the assistance provided by teachers and peers, and whether it accords with the attributes of assistance within the ZPD; that is, whether it represents scaffolded assistance. One form of assistance that has received much attention in recent L2 research has been assistance as feedback on learners’ errors in language use, termed corrective feedback (CF).

One contentious issue is which type of CF is most conducive to L2 learning (Ellis, 2009). For example, in the case of feedback provided on students’ oral production, there is disagreement about whether explicit feedback (e.g., direct correction) or implicit feedback (e.g., recasts, prompts) is more effective (see Long, 2007; Lyster, 2004). A similar debate is evident in the literature on written CF. Studies that have compared the effectiveness of direct

(i.e., providing reformulations) and indirect written CF (e.g., signalling the occurrence of an error via the use of symbols), have yielded no conclusive results about which form of feedback is the most effective (for a summary of these studies see Bitchener & Storch, 2016).

From a SCT perspective, there is no single or predetermined type of CF that is best for learning. Rather, for CF to be effective, it needs to take into consideration the learner's current and potential level of performance (i.e., their ZPD). The feedback provided needs to be dynamic, adjusted in terms of explicitness and specificity in response to the learner's signs of L2 development. Providing the same level of support to all learners, or to an individual learner regardless of the nature of their ability to take advantage of the feedback provided, may constrain L2 development. However, providing such carefully attuned feedback may be difficult to implement in a classroom because it implies individually tailored assistance. A number of studies (see the next section on empirical research) have investigated how scaffolded CF can be provided to L2 learners and the impact it has on L2 development.

Another issue of relevance to L2 instruction is the kind of tasks that provide optimal conditions for L2 learning. From a SCT perspective, tasks that encourage learners to use the tools at their disposal to mediate their performance are ideal. A number of studies have focused in particular on the kind of activities that can encourage languaging; that is, encourage learners to use language as a mediational tool in self-directed and other-directed talk. A related and contentious issue is L1 use in such languaging. Whereas L1 use is generally frowned upon in L2 classes, SCT views both the L1 and L2 as mediational resources. A number of studies (e.g., Centeno-Cortés & Jiménez-Jiménez, 2004; de Guerrero & Villamil, 2000; Storch & Aldosari, 2010) have investigated the extent of L1 use and the functions that the L1 serves when learners engage in languaging to determine whether some use of the L1 can be perceived as beneficial for L2 learning.

Key Concepts

ZPD: The difference between a learner's two levels of performance: performance with and without assistance. Performance with assistance is more indicative of the learner's potential for development.

Scaffolding: Finely attuned and dynamic assistance that is responsive to the learner's needs.

Mediation: The act of using tools to complete or enhance an action or a process (e.g., thinking process).

Tools: Artefacts created and used by humans. Tools can be material, such as books, or symbolic, such as language or gestures.

Languaging: Verbalisation of thinking processes in problem-solving activities. Languaging can be self-directed (private speech) or other-directed.

Empirical Evidence

ZPD, Scaffolding, and ISLA

As discussed earlier, from a SCT perspective, assistance is key to cognitive development, but only if that assistance is scaffolded; that is, attuned to the individual learner's potential abilities (ZPD) and contingently responsive to any changes in these abilities. Much has been written about scaffolding in mainstream education (e.g., Fernández, Wegerif, Mercer, &

Rojas-Drummond, 2015; Mercer, 1994; Moschkovich, 2015) and in computer-mediated instructional contexts (e.g., Girault & d'Ham, 2014), as well as of the challenges of providing such assistance to both L1 and L2 learners in whole-class activities. For example, the study by Hammond and Gibbons (2005), conducted in a number of high schools in Australia where ESL learners are taught content (e.g., science) and language concurrently, showed that it was only experienced teachers who were able to provide scaffolded assistance intuitively during teacher–student oral interactions. A study by Guk and Kellog (2007), which compared teacher-fronted activities and group work in an EFL primary class in Korea, found more evidence of scaffolding in the group interactions than in whole-class teacher–student interactions.

There are also challenges in ensuring that the CF provided to L2 learners on their writing is a form of scaffolded assistance. Providing such feedback orally in teacher–student conferences may provide opportunities for individualised scaffolded assistance. The study that is most often cited as illustrative of scaffolded CF on writing is that by Aljaafreh and Lantolf (1994). The authors operationalised scaffolded feedback as a “regulatory scale” with 12 levels of assistance, from the most implicit to the most specific and explicit. The small-scale study involved a tutor providing carefully scaffolded CF to three ESL learners in a series of oral one-on-one conferences. The nature of the CF the tutor provided depended on the learner’s response. The tutor began with the most implicit type of CF (e.g., inviting the learner to reread their text), and then, if necessary, the CF became more explicit (e.g., directing the learner to a specific sentence or providing the correct form). Aljaafreh and Lantolf reported that all the learners showed different developmental trajectories over time, depending on the grammatical structure that was targeted by the CF. However, it is important to note here that what was taken as evidence of development was not only the more accurate use of these structures in successive compositions, but also the learner’s need for less explicit forms of CF over time. Aljaafreh and Lantolf argue that if we look only at the accurate use of structures as evidence of development we may not fully capture a learner’s progress. Measures of development also need to consider if there has been any change in the quality of the assistance the learner requires. If a learner, for example, requires less explicit feedback over time in order to self-correct the use of a particular structure, this is also a sign of development. It implies a movement from other-regulation (reliance on the expert) to greater self-regulation. Proponents of dynamic assessment (e.g., Poehner & Lantolf, 2010) suggest that learners should be assessed on their performance with assistance in completing challenging tasks. The learner’s overall score is then composed of two scores: a score on task performance, which measures the learner’s current competence, and a score that reflects the nature of assistance the learner required to complete the task, which measures potential competence.

The regulatory scale developed by Aljaafreh and Lantolf in 1994 has since been used in a small number of studies (e.g., Erlam, Ellis, & Batstone, 2013; Nassaji & Swain, 2000). The findings of these studies confirm that carefully scaffolded CF may be more beneficial than randomly provided CF (Nassaji & Swain, 2000) or uniformly explicit CF (Erlam et al., 2013). However, providing such individualised feedback is very time-consuming (Erlam et al., 2013) and perhaps unrealistic in large classes.

Another way of implementing scaffolded feedback on L2 writing is to combine written CF and oral classroom sessions where scaffolded feedback is provided on certain targeted structures (e.g., the most common errors found in the learners’ writing). For example, Nassaji’s (2012) study compared the impact of three types of CF delivered in three consecutive weeks on two targeted structures (the use of English articles and prepositions). The first type

of feedback was nonscaffolded (explicit written reformulations). The second and third types involved oral classroom interactions between the teacher and students. In the second round of feedback, the feedback was minimally scaffolded (oral reformulations were provided only to learners who failed to self-correct). The third round of feedback involved a negotiation process: it began with prompts, which are considered (although not by all researchers) as a more implicit form of CF, and gradually became more explicit (i.e., reformulation), but only if needed. The study employed a pretest/posttest design, and thus after each type of feedback was delivered, the learners completed a posttest. Posttest results confirmed that scaffolded (negotiated) feedback resulted in greater accuracy scores compared to the other forms of CF but only for the use of articles. The negotiations that formed part of the scaffolded feedback also resulted in the learners gaining a greater understanding of articles, a rule-based structure (unlike prepositions).

Scaffolded CF on writing can also be implemented in L2 classes that adopt a process approach to writing instruction or in graduate programs where students submit multiple drafts of their writing. In such contexts, scaffolded CF would take the form of providing very implicit CF on early drafts (e.g., notes in the margin), becoming more explicit in successive drafts but only on structures that learners fail to self-correct. The study by Morton, Thompson, and Storch (2014) is one of the few studies that adopted a SCT perspective to retrospectively investigate the nature of the feedback provided by a supervisor (Storch) to her MA student on three drafts of a literature review chapter. The study found that most of the CF provided on all three drafts was explicit, provided in the form of deletions and reformulations. These findings suggested that the supervisor may have missed opportunities to provide scaffolded CF; that is, feedback attuned and responsive to the learner's capacities.

In the studies discussed thus far, the CF was provided by the teacher ('the expert') in oral or written form. Opportunities for scaffolded feedback may also be available from peers, fellow novices, working in small group/pair work either in peer response activities or collaborative writing activities (see discussion in the next section). In such activities learners provide each other with oral feedback but on written texts. For example, a series of descriptive studies on classroom peer response activities, undertaken by de Guerrero and Villamil with intermediate ESL learners in Puerto Rico, analysed the nature of feedback peers provided to each other (see de Guerrero & Villamil, 1994, 2000; Villamil & de Guerrero, 1996, 1998). The researchers reported that the learners provided effective assistance using a range of scaffolding strategies such as advising, requesting clarifications, and providing mini grammar lessons when needed. Furthermore, Brooks and Swain's (2009) small-scale study ($N = 4$), comparing the effects of peer and expert feedback, found that peer feedback was in fact more effective than expert feedback. The researchers suggested that the feedback provided by the expert dealt with structures that were perhaps beyond the learners' ZPD. Peers, on the other hand, provided each other with assistance that was more attuned to their own needs and developmental stages.

Using Language as a Meditational Tool in ISLA

Language, as a meditational tool, has been investigated in its two forms: self-directed talk and collaborative talk. The incidence of self-directed talk has been reported by a number of studies, including in whole-class activities, where students respond to the teacher's questions (e.g., Ohta, 2001), in complex problem-solving individual activities (e.g., Negueruela, 2008) and in pair activities (e.g., Watanabe, 2014). Self-directed talk includes self-repetitions,

sounding out different language forms to check if they sound correct, and self-directed questions. Such forms of talk enable individuals to engage in self-scaffolding (e.g., Knouzi, Swain, Lapkin, & Brooks, 2010; Negueruela, 2008; Swain et al., 2011; Watanabe, 2014). For example, Knouzi et al. (2010) asked French L2 learners to explain aloud their understanding of a series of sentences containing passive structures. The study reported that the learners used a range of strategies to scaffold their performance, such as drawing connections between new and prior knowledge, thinking of concrete examples, and self-assessing their understanding. Furthermore, the quality and quantity of languaging episodes were found to be correlated to the posttest results measuring knowledge of these passive structures.

There are various instructional strategies and tasks that may encourage learners to engage in self-directed forms of languaging. These strategies include asking learners to verbalise their thoughts, as in the study by Knouzi et al. discussed earlier. Another strategy is to ask students to write down their thoughts or reflections. A small number of studies have explored the impact of written forms of languaging (e.g., Ishikawa, 2013; Suzuki, 2012). In these studies, conducted with EFL learners in Japan, the written languaging was done in the learners' L1 (Japanese). In Ishikawa's study, the learners engaged in written languaging in two phases of a translation task. In Suzuki's (2012) study, the learners engaged in written reflections on receipt of teacher feedback on their writing. Specifically, the learners were asked to write explanations for why they thought their language forms were corrected. The study found that when participants understood why their language was corrected (as evidenced by the written languaging episodes) they were more likely to incorporate the corrections in their revised draft. Suzuki suggests that providing learners with opportunities to reflect about their linguistic knowledge facilitated the learners' L2 development.

An activity that has shown to elicit learners' languaging is collaborative writing tasks, where learners jointly co-author a text (see Storch, 2013 for an extended discussion of collaborative writing). Collaborative writing is a challenging task, much more so than writing individually. The production of a joint text means that the co-authors need to engage in negotiations about what ideas to include in their joint text and how to express their ideas. These negotiations can be time-consuming and not always easily resolved (DiNitto, 2000; Storch, 2013). However, such negotiations can also be a positive force, stimulating an exchange of ideas and exposure to new perspectives. Furthermore, collaborative writing provides a natural environment for peer feedback, as learners make suggestions and consider counter-suggestions about word choices and grammatical structures when co-constructing their text.

During collaborative writing tasks learners engage in self-directed and other-directed speech. However, because the self-directed speech occurs in the presence of another learner, the boundaries between self-directed and other-directed speech blur. Self-directed questions or expressions of uncertainty when vocalised may elicit a response from other learners in the small group or pair. This response can be in the form of, for example, a suggestion, an explanation, or a repair (see for example, Fernández Dobao, 2012, 2014; Storch, 2002).

These instances of languaging about specific linguistic items have been operationalised in the literature as language-related episodes (LREs) (Swain & Lapkin, 2002). LREs are occasions for language learning. A number of studies have indeed shown a positive relation between the quantity of languaging episodes (i.e., LREs) and language learning gains (e.g., Kim, 2008; Storch, 2002). Moreover, Fernández Dobao (2016) found that students' vocabulary learning benefitted from collaborative dialogue even when they did not actively participate in the dialogue but merely listened to their peers engaging in languaging episodes.

However, not all pair/group work is conducive to L2 learning. Storch (2002) identified four distinct relationships that intermediate adult ESL learners formed when working on a

range of language tasks: collaborative, expert/novice, dominant/dominant, and dominant/passive. The study found evidence of collective peer scaffolding and of learning gains predominantly in pairs that collaborated or formed an expert/novice pattern. Subsequent studies in face-to-face (e.g., Edstrom, 2015; Watanabe & Swain, 2007) and computer-mediated (e.g., Li & Zhu, 2013) peer work have confirmed the superiority of collaborative patterns of peer interaction for L2 learning. Students who collaborate are also more likely to enjoy the activity (Li & Zhu, 2013; Storch, 2004). Vygotsky (1978) argued that the affective and cognitive dimensions of learning cannot be separated.

Another activity that can promote languaging is asking pairs or small groups of students to consider the feedback provided by the teacher. A series of studies informed by Swain and Lapkin (2002) investigated learners' deliberations over feedback they received on their jointly written tasks, and the impact of these deliberations on revisions (e.g., Brooks & Swain, 2009; Storch & Wigglesworth, 2010a, 2010b; Tocalli-Beller & Swain, 2005). What these studies show is that these deliberations, where learners questioned, discussed, and explained language conventions, resulted in improved revisions. Learners were also more likely to remember the feedback that they deliberated about, and then to use this newly gained knowledge when they subsequently revised their original draft. In contrast, feedback that was merely accepted was less likely to be remembered and used in revisions (Storch & Wigglesworth, 2010a; Tocalli-Beller & Swain, 2005). However, these studies also found instances where the learners rejected the feedback because it violated earlier learnt language rules or was perceived to alter their intended meaning (Storch & Wigglesworth, 2010b; Swain & Lapkin, 2002).

What these findings remind us is that learners "need to be understood as people, which in turn means we need to appreciate their human agency. As agents learners actively engage in constructing the terms and conditions of their learning" (Lantolf & Pavlenko, 2001, p. 145). SCT views learners as active agents who assign relevance and significance to certain actions. For example, when receiving CF, learners exercise volitional control over what they notice in the feedback and whether they accept, question, or reject the feedback they receive.

L2 learners also have a choice of language to draw upon as a resource. In studies investigating pair and small group work on a range of language tasks, the L1 has been reported to be used as a tool to deliberate about language choice and form, and to gain a better understanding of challenging task requirements (see Azkarai & Garcia Mayo, 2015; de Guerrero & Villamil, 2000; Storch & Aldosari, 2010). For example, in a study of group work in a business subject, Yang (2014) found that learners used their shared L1 gainfully to solve complex mathematical problems and then generate a report written in the L2. Other researchers reported on how both L1 and L2 were used by learners in self-directed private speech when completing problem-solving tasks (e.g., Centeno-Cortés & Jiménez-Jiménez, 2004). What these studies suggest is that the judicious use of the L1 as a cognitive tool may support L2 learning and that policies that forbid any use of the L1 in L2 classes may deprive the students of an important cognitive tool.

Pedagogical Implications

What SCT implies for L2 learning/instruction is the need for two key ingredients: challenge and effective support. Effective L2 learning is more likely to occur when learners are presented with challenging tasks; that is beyond their current level of development. Challenging tasks will push learners to use their language (L2 and some L1) as cognitive tools to resolve any language related issue they encounter. Such challenging tasks, however, need to be

coupled with appropriate forms of support; namely scaffolded assistance. Scaffolded assistance will not only enable the learners to complete challenging tasks but also to complete similar tasks in the future independently. Research informed by SCT has shown that experienced teachers do provide such scaffolding (Gibbons, 2003; Hammond & Gibbons, 2005) as do peers, when engaging in collaborative pair and small group activities (e.g., Fernández Dobao, 2014; Storch, 2002). The other advantage of small group and pair work, if carefully designed and monitored, is that it also provides learners with opportunities to collectively scaffold their performance and to verbalise or language (Swain, 2006) their thinking process, in other- and self-directed talk. Such opportunities to language may be absent in teacher-directed classes (Guk & Kellog, 2007).

Teaching Tips

- Reflect on feedback practices: does the corrective feedback provided to learners on their writing show attributes of scaffolding?
- Tasks should be challenging and interesting so that learners are motivated and pushed to engage in languaging.
- Collaborative writing activities need to be carefully designed and monitored. Simply assigning students to write in pairs does not mean that they will work collaboratively.

Future Directions

Research to date on ISLA informed by SCT is still relatively small in volume and size. Many of the studies are small-scale case studies, not surprising perhaps given the ethnographic approach to data collection and the qualitative analysis of that data that these studies deploy. Furthermore, many studies have been conducted in ESL and EFL settings. Studies conducted with learners of languages other than English are still relatively rare (e.g., Fernández Dobao's 2012, 2014 studies conducted with learners of L2 Spanish). Clearly, more research in a diverse range of settings and student cohorts is needed. In the following I outline some research projects focusing on the two central constructs of ZPD (and scaffolding) and language as a mediating tool.

The construct of ZPD and scaffolded assistance, when provided in the form of CF on writing, has shown to be effective but time-consuming. Thus one possible solution is to provide scaffolded CF in computer-mediated form. Poehner and Lantolf (2010) describe a computerised form of dynamic assessment based on a predetermined scale of assistance consisting of a range of hints offered to learners for each test item. These hints range from implicit to explicit. For example, in the case of a language test, when a learner produces an incorrect response, the learner is provided in the first instance with an implicit hint (e.g., a suggestion to think about their response again). A second attempt that is also incorrect elicits a more explicit hint (e.g., a suggestion to think about the particular language structure). The final hint is the most explicit, providing the correct response along with an explanation. There are a number of advantages to such a computerised scaffolded CF delivery: the level of assistance provided is responsive to the learner's need (although not very finely tuned), the system can deal with a number of targeted structures, and the CF can be offered to large cohorts of students. However, using the computer to deliver CF on writing may have other, perhaps less desirable effects on learner engagement with the feedback, as suggested by a

critical review of research investigating the impact of automated assessment and feedback programs on L2 development (Stevenson & Phakiti, 2014). Thus scaffolded CF that is computer mediated is clearly an important area for future investigations.

Another suggested line of future research related to the construct of the ZPD is action research undertaken by L2 teachers. A number of studies have reported that supervisors (e.g., Basturkmen, East, & Bitchener, 2014; Morton et al., 2014) and L2 teachers (e.g., Al Shahrani & Storch, 2014) may be unaware of the amount and type of feedback they provide to their students. Weissberg (2006), among others, has called for teachers to critically reflect on their own feedback practices. The widespread use of interim drafts, in L2 writing classes and in graduate supervision, provides L2 teachers and supervisors working with L2 students with the opportunity to do so via retrospective action research. Using the ZPD as a theoretical framework, such retrospective action research could investigate whether indeed the kind of CF provided to learners accords with the traits of scaffolded feedback. This kind of investigation involves not only an analysis of the nature and focus of the feedback we provide to L2 learners on successive drafts (see Morton et al., 2014) or on different texts produced over time, but also a concurrent investigation of learners' writing to examine whether the feedback offered is responsive to the L2 learners' changing needs. Such research could help transform teachers' written CF practices. Subsequent investigations could examine whether the implementation of scaffolded CF encourages learners to become increasing self-regulated, able to self-correct.

A strategy that may enhance the outcomes of CF on writing is to engage students in deliberations about the feedback they receive. These deliberations are a form of languaging—using language to notice and acknowledge the feedback, to understand or to question why the feedback was given. A number of studies (e.g., Storch & Wigglesworth, 2010a, 2010b; Suzuki, 2012) provide evidence that such languaging is associated with improved accuracy of revised texts. However, what most of these studies lack is evidence showing that learners can use this knowledge as a resource in independent performance on new writing. Thus future research needs to investigate the impact of languaging in deliberations on CF and on language learning.

Given the widespread use of online collaborative writing platforms such as wikis or Google Docs in L2 writing classes (see review in Storch, 2013), another important area of investigation is the impact that computer-mediated forms of interaction have on languaging. Rouhshad and Storch (2016) is one of the few studies to date that has compared the nature of languaging in face-to-face and computer-mediated forms of interaction (using Google Docs) when learners completed collaborative writing tasks. The study found that the tool (i.e., computer) impacted not only on the relationships learners formed when working in pairs but also on the nature of languaging. In the computer-mediated interactions, collaboration was rare, as were instances of languaging (measured via LREs). These studies confirm the findings of other studies that compared the opportunities for language learning in face-to-face and computer-mediated forms of interaction on oral tasks (e.g., Loewen & Wolff, 2016). Future investigations need to further explore not only the impact of this new tool on interaction and on the nature of languaging, but also ultimately on language learning.

References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *Modern Language Journal*, 78, 465–483.
- Al Shahrani, A., & Storch, N. (2014). Investigating teachers' written corrective feedback practices in a Saudi EFL context: How do they align with their beliefs, institutional guidelines, and students' preferences? *Australian Review of Applied Linguistics*, 37, 101–122.

- Azkarai, A., & Garcia Mayo, M. P. (2015). Task-modality and L1 use in EFL oral interaction. *Language Teaching Research*, 19(5), 550–571.
- Basturkmen, H., East, M., & Bitchener, J. (2014). Supervisors' on-script feedback comments on drafts of dissertations: Socialising students into the academic discourse community. *Teaching in Higher Education*, 19(4), 432–445.
- Bitchener, J., & Storch, N. (2016). *Written CF for L2 development*. Bristol: Multilingual Matters.
- Brooks, L., & Swain, M. (2009). Languageing in collaborative writing: Creation and response to expertise. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 58–89). New York: Routledge.
- Centeno-Cortés, B., & Jiménez-Jiménez, A. (2004). Problem-solving tasks in a foreign language: The importance of the L1 in private verbal thinking. *International Journal of Applied Linguistics*, 14(1), 7–35.
- de Guerrero, M.C.M., & Villamil, O. S. (1994). Social-cognitive dimensions of interaction in L2 peer revision. *Modern Language Journal*, 78(4), 484–496.
- de Guerrero, M.C.M., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *Modern Language Journal*, 84(1), 51–68.
- DiNitto, R. (2000). Can collaboration be unsuccessful? A sociocultural analysis of classroom setting and Japanese L2 performance in group tasks. *Journal of the Association of Teachers of Japanese*, 34(2), 179–210.
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Norwood, NJ: Ablex.
- Edstrom, A. (2015). Triads in the L2 classroom: Interaction patterns and engagement during a collaborative task. *System*, 52, 26–37.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107.
- Erlam, R., Ellis, R., & Batstone, R. (2013). Oral corrective feedback on L2 writing: Two approaches compared. *System*, 41, 257–268.
- Fernández, M., Wegerif, R., Mercer, N., & Rojas-Drummond, S. (2015). Re-conceptualizing “scaffolding” and the zone of proximal development in the context of symmetrical collaborative learning. *Journal of Classroom Interaction*, 50(1), 54–72.
- Fernández Dobao, A. (2012). Collaborative writing tasks in the L2 classroom: Comparing group, pair, and individual work. *Journal of Second Language Writing*, 21(1), 40–58.
- Fernández Dobao, A. (2014). Attention to form in collaborative writing tasks: Comparing pair and small group interaction. *Canadian Modern Language Review*, 70(2), 158–187.
- Fernández Dobao, A. (2016). Peer interaction and learning: A focus on the silent learner. In M. Sato & S. Ballinger (Eds.), *Peer interaction in second language learning: Pedagogical potential and research agenda* (pp. 33–61). Amsterdam/Philadelphia: John Benjamins.
- Frawley, W., & Lantolf, J. (1985). Second language discourse: A Vygotskian perspective. *Applied Linguistics*, 6(1), 19–44.
- Gibbons, P. (2003). Teacher interactions with ESL students in content-based classroom. *TESOL Quarterly*, 37(2), 247–273.
- Girault, I., & d'Ham, C. (2014). Scaffolding a complex task of experimental design in chemistry with a computer environment. *Journal of Science, Education and Technology*, 23(4), 514–526.
- Gregg, K. R. (1993). Taking explanations seriously; or: Let a couple of flowers bloom. *Applied Linguistics*, 14(3), 276–294.
- Guk, I., & Kellog, D. (2007). The ZPD and whole class teaching: Teacher-led and student-led interactional mediation of tasks. *Language Teaching Research*, 11(3), 281–299.
- Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect*, 20(1), 6–30.
- Ishikawa, M. (2013). Examining the effect of written languageing: The role of metanotes as a mediator of second language learning. *Language Awareness*, 22(3), 220–233.

- Kim, Y. J. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *Modern Language Journal*, 92(1), 114–130.
- Knouzi, I., Swain, M., Lapkin, S., & Brooks, L. (2010). Self-scaffolding mediated by language: Microgenetic analysis of high and low performers. *International Journal of Applied Linguistics*, 20, 23–49.
- Lantolf, J. P. (2000). Second language learning as a mediated process. *Language Teaching*, 33(1), 79–96.
- Lantolf, J. P., & Pavlenko, A. (2001). Second language activity theory: Understanding second language learners as people. In M. Breen (Ed.), *Learner contributions to language learning* (pp. 141–158). London: Longman.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of L2 development*. Oxford: Oxford University Press.
- Leontiev, A. N. (1978). *Activity, consciousness, and personality*. Englewood Cliffs, NJ: Prentice Hall.
- Li, M., & Zhu, W. (2013). Patterns of computer-mediated interaction in small writing groups using wikis. *Computer Assisted Language Learning*, 26(1), 61–82.
- Loewen, S., & Wolff, D. (2016). Peer interaction in F2F and CMC contexts. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 163–184). Mahwah, NJ: John Benjamins.
- Long, M. H. (1985). Input and second language acquisition theory. In S. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 377–393). Rowley, MA: Newbury House.
- Long, M. H. (1990). The least a second language acquisition theory needs to explain. *TESOL Quarterly*, 24(4), 649–666.
- Long, M. H. (2007). *Problems in SLA*. New York: Lawrence Erlbaum.
- Luria, R. (1973). *The working brain*. New York: Basic Books.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Mercer, N. (1994). Neo-Vygotskian theory and classroom education. In B. Steirer & J. Maybin (Eds.), *Language, literacy and learning in educational practice* (pp. 92–110). Clevedon: Multilingual Matters.
- Morton, J., Thompson, C., & Storch, N. (2014). Feedback in the supervision of postgraduate students: Insights from the work of Vygotsky and Bakhtin. *Journal of Academic Language and Learning*, 8(1), 24–36.
- Moschkovich, J. (2015). Scaffolding student participation in mathematical practices. *International Journal of Mathematics Education*, 47(7), 1067–1078.
- Nassaji, H. (2012). Correcting students' written grammatical errors: The effects of negotiated versus nonnegotiated feedback. *Studies in Second Language Learning and Teaching*, 1, 315–334.
- Nassaji, H., & Swain, M. (2000). A Vygotskian perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness*, 9(1), 34–51.
- Negueruela, E. (2008). Revolutionary pedagogies: Learning that leads to second language development. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 189–227). London: Equinox.
- Ohta, A. S. (2001). *Second language acquisition processes in the classroom: Learning Japanese*. Mahwah, NJ: Lawrence Erlbaum.
- Piaget, J. (1977). *The development of thought: Equilibrium of cognitive structures*. New York: Viking.
- Poehner, M. E. (2008). Both sides of the conversation. In J. P. Lantolf & M. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 33–56). London: Equinox.
- Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity*, 17, 312–330.
- Roth, W. M., & Radford, L. (2010). Re/thinking the zone of proximal development (systematically). *Mind Culture and Activity*, 17(4), 299–307.

- Rouhshad, A., & Storch, N. (2016). A focus on mode: Patterns of interaction in face-to-face and computer-mediated contexts. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 267–289). Amsterdam/Philadelphia: John Benjamins.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(1), 4–13.
- Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19(1), 51–65.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52, 119–158.
- Storch, N. (2004). Using activity theory to explain differences in patterns of dyadic interactions in an ESL class. *Canadian Modern Language Review*, 60, 457–480.
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol: Multilingual Matters.
- Storch, N., & Aldosari, A. (2010). Learners' use of first language (Arabic) in pair work in an EFL class. *Language Teaching Research*, 14, 355–375.
- Storch, N., & Wigglesworth, G. (2010a). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, 32, 303–334.
- Storch, N., & Wigglesworth, G. (2010b). Students' engagement with feedback on writing: The role of learner agency/beliefs. In R. Batstone (Ed.), *Sociocognitive perspectives on language use and language learning* (pp. 166–185). Oxford: Oxford University Press.
- Suzuki, W. (2012). Written languaging, direct correction, and second language writing revision. *Language Learning*, 62, 1110–1133.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford: Oxford University Press.
- Swain, M. (2006). Languaging, agency and collaboration in advanced language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). London: Continuum.
- Swain, M., Kinnear, P., & Steinman, L. (2011). *Sociocultural theory in second language education: An introduction through narratives*. Bristol: Multilingual Matters.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37, 285–304.
- Tocalli-Beller, A., & Swain, M. (2005). Reformulation: The cognitive conflict and L2 learning it generates. *International Journal of Applied Linguistics*, 15(1), 5–28.
- van Compernelle, R. A., & Williams, L. (2013). Sociocultural theory and second language pedagogy. *Language Teaching Research*, 17(3), 277–281.
- van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural approach*. Amsterdam: Kluwer Academic.
- Villamil, O. S., & de Guerrero, M. C. M. (1996). Peer revision in the L2 classroom: Socio-cognitive activities, mediating strategies, and aspects of social behavior. *Journal of Second Language Writing*, 5, 51–75.
- Villamil, O. S., & de Guerrero, M. C. M. (1998). Assessing the impact of peer revision on L2 writing. *Applied Linguistics*, 19, 491–514.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), *The concept of activity in Soviet psychology* (pp. 144–188). Armonk, NY: M. E. Sharpe.
- Vygotsky, L., & Luria, A. (1994). Tool and symbol in child development. In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 99–174). Oxford: Blackwell.
- Watanabe, Y. (2014). *Collaborative and independent writing: Japanese university English learners' processes, texts and opinions* (Unpublished PhD thesis). Ontario Institute of Studies in Education, University of Toronto.

- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11(2), 121–142.
- Weissberg, R. (2006). Scaffolded feedback: Tutorial conversations with advanced L2 writers. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 246–265). New York: Cambridge University Press.
- Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.
- Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.
- Wilson, K., & Devereux, L. (2014). Scaffolding theory: High challenge, high support in Academic Language and Learning (ALL) contexts. *Journal of Academic Language and Learning*, 8, A91–A100.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry*, 17, 89–100.
- Yang, L. (2014). Examining the mediational means in collaborative writing: Case studies in undergraduate ESL students in business courses. *Journal of Second Language Writing*, 23(1), 74–89.



Taylor & Francis

Taylor & Francis Group
<http://taylorandfrancis.com>

Section II

Approaches to Second Language Instruction



Taylor & Francis

Taylor & Francis Group
<http://taylorandfrancis.com>

6

Content-Based Language Teaching

Roy Lyster

Background

Content-based language teaching (CBLT) is an instructional approach in which nonlinguistic curricular content such as geography or science is taught to students through the medium of a language that they are learning as an additional language. CBLT comes in many different shapes and sizes, and in fact is called by other names and acronyms, including content-based instruction (CBI) and content and language integrated learning (CLIL). Whether called CBLT, CBI, or CLIL, a range of instructional initiatives can be identified along a continuum with *language-driven* programs at one end and *content-driven* programs at the other end (see Figure 6.1).

At the language-driven end of the spectrum are foreign language classes that promote target language development by incorporating a focus on theme-based content but without high-stakes assessment of students' content knowledge. The goal of such classes is to "help learners develop their L2 competence within specific topic areas" (Brinton, Snow, & Wesche, 2003, p. 19). Another anticipated goal is to transform foreign language classrooms "into sites where intellectually stimulating explorations can become the norm rather than the exception" (Cammarata, 2016, p. viii).

Toward the middle of the continuum are program models in which students study one or two subjects in the target language, usually in tandem with a foreign language or language arts class. This is the model adopted by many CLIL programs in Europe and elsewhere (Coyle, Hood, & Marsh, 2010). There are many varieties of CLIL, but it typically begins in secondary schools and usually offers less than half the curriculum in the target language (often one content course and one EFL course). Important to mention is that CLIL is used by many as "a generic umbrella term for bilingual, content-based education" (Ruiz de Zarobe, 2008, p. 61) in the same way that CBLT is being used in this chapter.

Also in the middle of the continuum is the English-medium CBLT program implemented in China that teaches content areas that are not part of the formal curriculum such as "nature and society" and "science and life" usually for two lessons per week at the middle school level (Hoare, 2010). An example at the postsecondary level is the University of Ottawa's adjunct model of CBLT, which enables nonfrancophone students to take

Roy Lyster

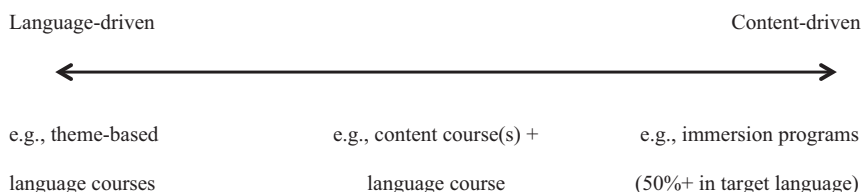


Figure 6.1 Range of CBLT settings

regular content courses offered in French with francophone students for whom the content courses were originally designed. At the same time, they are “sheltered as a group in a separate credit language course related to the content course” (Burger & Chrétien, 2001, p. 85). Also at the postsecondary level, language and literature departments often offer upper-level undergraduate content courses in the target language, such as Italian social and physical geography courses taught in Italian (Rodgers, 2006) or courses “structured around literary and/or cultural themes in the Francophone, Spanish, or Latin American world” (Rodgers, 2015, p. 116). Content-based EFL courses have also been introduced at postsecondary levels in Japan, where a task-based approach to CBLT has shown considerable promise for teaching courses in comparative culture (Lingley, 2006). Noteworthy in this regard is a recent special issue of *System* (Vol. 54) exploring the interface between task-based language teaching and CBLT, based on the premise that CBLT aims to achieve its goal of integrating language and content “by means of tasks that are cognitively engaging for the learners” (García Mayo, 2015, p. 1).

At the content-driven end of the spectrum are school-based language immersion programs that aim for additive bilingualism by providing a substantial portion of students' subject-matter instruction through the medium of a language that they are learning as a second, foreign, heritage, or indigenous language. The remaining proportion of the curriculum is provided through the medium of a shared primary language, which normally has majority status in the community. At the elementary level, at least 50% of the curriculum is taught in the immersion language, whereas continuation programs at the secondary level include a minimum of two subject courses in the immersion language. Important to mention is that some European bilingual programs meeting these immersion criteria are designated as CLIL programs in cases where the target language is English (see Llinares & Dafouz, 2010, regarding such programs in Madrid).

Immersion programs have been adopted in some countries to promote the learning of a second co-official language. Examples of these include French immersion in Canada (Lyster, 2007), Swedish immersion in Finland (Björklund, Mård-Miettinen, & Savijärvi, 2013), Catalan immersion in Spain (Arnau & Vila, 2013), Basque immersion in Spain (Cenoz, 2008), and Irish immersion in Ireland (Ó Baoill, 2007). In still other contexts, school-based CBLT programs have been designed to deliver at least half the curriculum through the medium of a regional language such as Breton and Occitan in France (Costa & Lyster, 2011) or an indigenous language such as Maori in New Zealand (Reedy, 2000) and, in the US, Hawaiian (Luning & Yamauchi, 2010) and Cherokee (Peter, 2014). Also in the US are a growing number of two-way immersion programs, which normally integrate a similar number of children from two different mother-tongue backgrounds (e.g., Spanish/English) and provide curricular instruction in both languages (Lindholm-Leary, 2001; for a recent review of research on both one-way and two-way immersion programs in the US, see Tedick & Wesely, 2015).

English as an international language is the target language of a variety of CBLT programs ranging from early immersion in Japan (Bostwick, 2001) and Brazil (French, 2007) to late immersion in Hong Kong (Hoare & Kong, 2008), as well as international schools such as the one described by Spezzini (2005) in Paraguay.

Often thought of as an extension of CLIL programs, English as a medium of instruction (EMI) in higher education is another rapidly expanding area of content-based instruction, especially in the European context (Coleman, 2006). EMI has grown as a result of the Erasmus program (also known as the European Region Action Scheme for the Mobility of University Students), which since its inception has enabled almost three million students to complete a part of their studies abroad (Feye & Krzaklewska, 2013). To facilitate mobility and to attract more international students, universities are increasingly offering programs in English rather than in the national language, thus enabling not only Erasmus participants but also local students to study content through a language other than their first language (L1). This is sometimes referred to as Integrating Content and Language in Higher Education (ICLHE), but “with language learning remaining of secondary importance” (Smit & Dafouz, 2012, p. 3). See Arias and Izquierdo (2015) for a similar description of EMI in Mexican higher education.

Yet another context for CBLT at the content-driven end of the continuum includes schools where minority-language students, typically whose parents have immigrated to the host society, find themselves without any L1 support and with a majority of native speakers of the target language. These are regarded as mainstream or even “submersion” classrooms. In many contexts, these students are left to their own devices to deal with the home/school language switch, as noted by Nicholas and Lightbown (2008): “characteristics of appropriate L2 instruction are often absent as learners are expected to learn the language and the school subject matter at the same time—more or less by ‘osmosis’” (p. 45). In many US schools, however, content-based ESL and “sheltered instruction” programs are available to better address the needs of minority-language students who are learning English while also learning curricular content through English. In content-based ESL, “teachers seek to develop the students’ English language proficiency by incorporating information from the subject areas that students are likely to study,” and sheltered instruction entails content courses for ESL learners taught normally by content (rather than ESL) specialists (Echevarría, Vogt, & Short, 2008, p. 13). These contexts have given rise to a useful teacher development tool known as the Sheltered Instruction Observation Protocol (SIOP; see Echevarría et al., 2008), which provides teachers with guidance in implementing subject area curriculum to students learning through a language other than their L1 while maintaining grade-level objectives. It includes techniques that make the content material accessible and that develop literacy skills as well as skills specific to L2 learners.

CBLT thus crosses a wide range of international contexts and instructional settings, including elementary, secondary, and postsecondary institutions. In spite of the tremendous differences across these contexts (some including majority-language and others minority-language students), there are some common pedagogical issues that arise at the interface of language and content teaching. As Wesche (2001, p. 1) argued, “the contexts have much in common, each involving learners struggling to master academic concepts and skills through a language in which they have limited proficiency, while at the same time striving to improve that proficiency.” She suggested that learners’ efforts in this endeavor “can be facilitated by considerably good teaching.” The next section will identify some of the pedagogical challenges that arise at the interface of content and language teaching. Drawing on empirical research, best practices to address the challenges will then be outlined.

Roy Lyster

Current Issues

One of the most widely substantiated outcomes of French immersion programs is that students' L1 development and achievement in subjects taught in French are similar to (or better than) those of nonimmersion students. Genesee (2004) confirmed that these findings related to L1 development and academic achievement in the L2 "have been replicated, for the most part, in other regions of the world where similar programs with majority language students have been implemented" (p. 551). Examples of such programs include the English immersion program at Katoh Gakuen in Japan (Bostwick, 2001) and the Swedish immersion program in Finland (Björklund et al., 2013).

Another finding that is common across French immersion programs is that students develop much higher levels of proficiency in French than do nonimmersion students studying French as a regular school subject (e.g., 40 min/day). Wesche and Skehan (2002, p. 227) attributed the overall positive outcomes of CBLT to its potential to "provide the motivating purpose for language learning, a naturalistic learning context that includes social and other pragmatic dimensions, and the possibility of form-focused activity." They concluded that, "together, these perhaps offer as close to a comprehensive environment for second language development as is possible in the classroom." They also cautioned, however, that CBLT is "not a panacea that can achieve success whatever the circumstances." They argued that, for CBLT to be effective, "it has to be carefully introduced and implemented and requires appropriate teacher training and adaption to local conditions."

Arguably related to these words of caution concerning the need for careful implementation and ongoing professional development is the finding that the L2 proficiency of French immersion students in Canada is good in some domains but not others. Specifically, French immersion students develop high levels of communicative ability but lower-than-expected levels of productive abilities with respect to grammatical accuracy, lexical variety, and sociolinguistic appropriateness (Harley, Cummins, Swain, & Allen, 1990). On a positive note, there is growing consensus that higher levels of proficiency will be attainable through improved instructional strategies.

Based on the outcomes of French immersion programs, Swain (1988) proposed that content teaching on its own is not necessarily good language teaching and needs to be manipulated and complemented in ways that maximize target language learning. Otherwise, she argued, use of the target language to teach content has limitations in terms of the range of the language forms and functions to which it exposes students. A powerful example of this pertains to the distribution of verb tenses used by French immersion teachers: 74%–75% in the present tense or imperative forms and only 14%–15% in the past tense (Harley, Allen, Cummins, & Swain, 1987; Lyster, 2007). The disproportionate use of present tense and imperative forms helps to explain gaps in French immersion students' L2 development, especially their limited use of conditional forms and their inaccurate use of past tense forms. Similarly and more recently, in their analysis of the oral production of Cherokee immersion students, Peter, Hirata-Edds, and Montgomery-Anderson (2008) observed a predominance of verbs in the imperative form in obligatory contexts for the present continuous. They concluded that the students' overuse of imperative forms was likely due to the fact that imperative forms were the verb forms used most frequently by teachers to address students.

The "functionally restricted" input to which immersion students are exposed (Swain, 1988, p. 74) has also been invoked to explain other gaps in French immersion students' language development. For example, their choice of second person pronouns—characterized by overuse of informal *tu* and underuse of formal *vous*—has been linked to the

absence of formal *vous* in classroom discourse (Swain, 1988) but also to teachers' use of *tu* to indicate indefinite reference and even plural reference as they address the whole class while expressing a sense of closeness with each individual (Lyster & Rebuffot, 2002). With respect to lexical clues available in teacher discourse to mark grammatical gender (another well-documented problem for immersion students), Poirier and Lyster (2014) reported that only half of the determiners and adjectives used by French immersion teachers and less than a third of all direct object third person clitic pronouns were clearly marked for grammatical gender. Finally, with respect to gaps in immersion students' sociolinguistic competence, Mougeon, Nadasdi, and Rehner (2010) reported that students' underuse of vernacular and other informal variants on the one hand, and their overuse of formal variants on the other, reflected their teachers' excessive use of formal variants at the expense of informal variants.

Another concern about content teaching on its own is that it can take on a lecture format without providing sufficient opportunities for interaction and student production. For example, Moriyoshi (2010) conducted an observational study of two postsecondary CBLT classes in Japan: a geography class and a sociology class taught in English. In addition to the analysis of 7.5 hours of video-recorded observations, the 76 participating students were administered questionnaires and the two native English-speaking teachers were interviewed and also completed a questionnaire. The results revealed that the instructors provided extensive comprehensible input to students, focusing exclusively on content, especially on vocabulary, while students had limited opportunities to produce the language. Of the total words spoken, the instructors uttered 93% and students the remaining 7%. Notwithstanding, both teachers and students perceived the CBLT classes in a positive light, considering them to be effective for improving both listening skills and content knowledge.

A study of English-medium mathematics and science classes in Malaysian high schools (Tan, 2011) stands out as a cautionary example of the issues that can arise when CBLT programs are adopted to teach through the medium of English as an international language before content and language teachers are adequately prepared for the major overhaul of instructional practices engendered by such a policy change. The study illustrates how collaboration between content and language teachers was thwarted not only by constraints such as exam-driven curricula and minimal training in CBLT but also by the expectation that math and science teachers would seek language support beyond class time from EFL teachers in the same school. Moreover, the content teachers perceived themselves as *only* content teachers and the EFL teachers perceived themselves as *only* language teachers.

The belief that one is either a content teacher or a language teacher has been noted across a wide spectrum of CBLT contexts. A French immersion teacher of Grade 1 was reported as saying, "From 9:00 until 3:30, I do not teach French. I teach subject matter, and French is learned through this content" (Salomone, 1992, p. 22). At the secondary level, an economics teacher in an EMI program in Hong Kong stated, "As a teacher of economics, I don't think it's necessary to have to teach them language at all. I won't" (Trent, 2010, p. 117), while lecturers teaching physics in English at Swedish universities claimed, "I don't teach language I teach physics" (Airey, 2012, p. 74). These beliefs, however, run counter to the educational axiom that all teachers are language teachers, which is a core tenet of CBLT.

Through interviews with four content teachers and four ESL teachers, Trent (2010) explored the prospects of collaborative relationships between language and content teachers in English-medium secondary schools in Hong Kong. He found rigid divisions between departments and a hierarchy of disciplines, with language being perceived at the

Roy Lyster

lower end. He proposed a set of solutions to move schools toward a more collaborative mindset. First, school personnel need to explore commonalities of what it means to be a “teacher” as opposed to the more specific identities of an “economics teacher” or a “language teacher.” Second, teachers need to be empowered to move away from traditional ways of working within independent departments toward the development of cross-curricular relationships. In a similar vein, his third proposal is for individual teachers and individual academic departments to develop a school-wide set of curriculum goals and a “whole-school identity.”

To make CBLT more language-rich and discourse-rich, several proposals have been made (see the following Teaching Tips). The next section reports on some of the pedagogical implications of a more robust and seamless integration of language and content in CBLT.

Teaching Tips

Recommendations for Ensuring a Language Focus in CBLT

- Draw students’ attention to specific form/meaning mappings by creating contrived contexts that allow students to notice L2 features in their full functional range (Swain, 1988).
- Engage students in carefully planned and guided communicative practice activities that focus their attention on potential problems and elicit particular uses of language (Allen, Swain, Harley, & Cummins, 1990).
- Emphasize academic language functions such as describing, explaining, and predicting (Dalton-Puffer, 2007).
- Adopt a counterbalanced approach that gives content and language objectives complementary status and that shifts students’ attention between language and content (Lyster, 2007).
- Present subject matter through knowledge relationships such as cause–effect, hypothesis, and comparison (Kong, 2009).
- Foster technical academic knowledge rather than only commonsense knowledge and build on students’ knowledge while pushing them to elaborate their ideas more fully (Kong & Hoare, 2011).
- Highlight the ways in which linguistic features of disciplinary-specific language construe particular kinds of meanings (Llinares, Morton, & Whittaker, 2012).

Empirical Evidence and Pedagogical Implications

A useful way for teachers to manage the integration of language and content is to adopt a counterbalanced approach to CBLT that shifts students’ attention between language and content, specifically toward language if the classroom is primarily content-driven, as is often the case in immersion classrooms, or toward content if the overall classroom context is predominantly language-driven, as with many foreign language classrooms.

A counterbalanced approach to CBLT can be operationalized as either reactive or proactive, but, for optimal effectiveness, both are best implemented in tandem. A reactive approach includes scaffolding techniques such as questions and feedback in response to students’ language production that serve to support student participation while ensuring that classroom interaction is a key source of learning. A proactive approach includes preplanned activities that draw students’ attention to language features that might otherwise not be used or even

noticed in classrooms focusing on content. The implementation of reactive and proactive approaches as complementary approaches to CBLT is in line with Day and Shapson's (1996) case studies of French immersion teachers that led to their conclusion that both planned language instruction and the many unplanned opportunities teachers can seize on to enhance language learning are of equal importance.

Key Concepts

Principles of a Counterbalanced Approach to CBLT

- Content and language objectives are interdependent.
- Shifting students' attention between content and language increases depth of processing and thus strengthens their metalinguistic awareness.
- Metalinguistic awareness is essential in CBLT because it serves as a tool for detecting linguistic patterns in content-based input to support continued language growth.
- Both preplanned instruction and unplanned opportunities for teachable moments are complementary and of equal importance.

A Reactive Approach to CBLT

Teachers can integrate content and language in seemingly spontaneous ways through a reactive approach (Lyster, 2007, 2016). Ostensibly unplanned opportunities can take the form of (1) teacher questions intended to increase both the quantity and quality of student output and (2) corrective feedback that serves to negotiate both form and meaning. Questioning and feedback techniques together provide learners with the scaffolding they need in order to understand, participate, and engage with both language and content.

A reactive approach is considered to encompass opportunities that are “seemingly spontaneous” and “ostensibly unplanned” because oral interaction may indeed seem spontaneous and unplanned. However, oral interaction is unlikely to reach its full potential as a key source of learning in CBLT unless teachers reflect on its many facets and then plan accordingly, using interactional strategies considered to create optimal conditions for learning.

Scaffolding techniques are at the core of CBLT and are requisite for students' academic success. The notion that learners can and should engage with language just ahead of their current level of ability is an essential part of CBLT. By means of the scaffolding provided by teachers, students are able to engage with content in a language they know only partially, because they can draw on the contextual clues provided in the scaffolding while also drawing on prior knowledge. One type of scaffolding assists students in understanding content presented through their L2 and another type supports them in productively using the L2 to engage with the content.

Teachers have at their disposal a wide range of instructional strategies that facilitate students' comprehension of curricular content through the target language. These include scaffolding techniques that give students many chances to understand the target language and curricular content. For example, teachers can build redundancy into their speech by using self-repetition and paraphrase, as well as multiple examples, definitions, and synonyms. In tandem with their verbal input, teachers can use props, graphs, and other graphic organizers (see Early, 2001; Mohan, 1986), as well as visual and multimedia resources. To further

Roy Lyster

facilitate comprehension, teachers can rely on extensive body language, including gestures and facial expressions, and a range of paralinguistic elements.

Scaffolding the interaction to facilitate comprehension, however, needs to be seen as a temporary support so that students progressively develop more advanced comprehension strategies that enable them to process the target language autonomously without the scaffolding. Instructional techniques that rely too much on linguistic redundancy, gestures, and other visual and nonlinguistic support are unlikely over time to make the kinds of increasing demands on the learners' language system that are necessary for continued L2 learning. This means that teachers need to engage in a delicate balancing act of providing, on the one hand, just the right amount of support to make the target language comprehensible, while being demanding enough, on the other hand, to ensure that learners engage in higher order cognitive skills.

Teachers need also to provide support for their students to use the target language productively. First, in their own interaction with students, teachers need to give students appropriate "wait time" to interpret questions and formulate responses. Second, they need to create many opportunities for students to use the target language, including role plays, simulations, debates, and presentations, while also using a variety of interactive groupings such as dyads, think-pair-share, and learning centres, in order to promote learning from and with peers (e.g., peer editing, peer tutoring, peer correction).

By providing the amount of assistance that students need until they are able to function independently, teachers can promote both language development and the acquisition of content knowledge. The image of the teacher scaffolding learners so they can express what they would be unable to express on their own provides a helpful metaphor for appreciating the strategic role played by teacher questions in CBLT, which are addressed next.

Teacher Questions

In their seminal study of classroom discourse, Sinclair and Coulthard (1975) found that the most typical teaching exchange consists of three moves: an initiating (I) move by the teacher; a responding (R) move by the student; and a feedback (F) move by the teacher.

The IRF sequence is seen as the quintessence of transmission models of teaching and typical of teacher-centred classrooms. It has been criticized for engaging students only minimally and for maintaining unequal power relationships between teachers and students. Nevertheless, the IRF sequence continues to permeate classroom discourse, probably because it helps teachers to monitor students' knowledge and understanding (Mercer, 1999). By assessing their students in an ongoing manner in the course of interaction, teachers are better equipped to plan and evaluate CBLT. Moreover, IRF exchanges can develop into more equal dialogue if, in the third turn, the teacher avoids evaluation and instead requests justifications or counterarguments (Nassaji & Wells, 2000). In this regard, the feedback move is more aptly seen as a follow-up move that aims to: (1) elaborate on the student's response or provide clarification; (2) request further elaboration, justification, explanation, or exemplification; and (3) challenge students' views (Haneda, 2005). Echevarria and Graves (1998) identified helpful questioning techniques designed to push students to elaborate on their answers in this way. This kind of push helps students to deepen their understanding of ideas and concepts and provides opportunities for students to use language that is more complex than that found in the shorter answers that are more typical of CBLT discourse.

Teaching Tips

Use Effective Follow-Up Questions
(Echevarria & Graves, 1998, pp. 162–164)

- “Tell me more about . . .”
- “What do you mean by . . .”
- “In other words . . .”
- “Why do you think that?”
- “How do you know?”
- “What makes you think that?”
- “Look at the page and tell me what you think the chapter will be about.”
- “What can you learn from reading this label?”
- “How are these plants different?”
- “Why would the colonists do that?”
- “Tell me more about that.”
- “On what basis would you group these objects?”
- “Why might that be?”
- “What makes you think this might be different?”

Corrective Feedback

Theoretical perspectives that run the gamut from skill acquisition theory to cognitive-interactionist and sociocultural orientations posit that corrective feedback (CF) is not only beneficial but may also be necessary for moving learners forward in their L2 development. Although empirical research has consistently demonstrated that the provision of CF is more effective than no CF, there are still many variables that interact to moderate its effectiveness. One of the variables specific to CBLT is the tension that arises when curricular objectives emphasize both content knowledge and L2 development.

The way in which teachers interact with their students is considered to be central to CBLT. In particular, CF provided during teacher–student interaction is one way for teachers to integrate a focus on language into their instructional practices. In contexts of CBLT, CF is generally considered to comprise recasts, explicit correction, and prompts. A recast is “the teacher’s reformulation of all or part of a student’s utterance, minus the error” (Lyster & Ranta, 1997, p. 46). Explicit correction also provides the correct form but, unlike recasts, “clearly indicates that what the student had said was incorrect” (p. 46). In contrast, prompts withhold correct forms and instead provide clues to prompt students to retrieve these forms from their current knowledge.

Llinares and Lyster (2014) analyzed oral interaction in nine Grade 4–5 classrooms that included (1) two CLIL classrooms in Spain with English as the target language, (2) four French immersion classrooms in Quebec, and (3) three Japanese immersion classrooms in the US. The comparison revealed that teachers in all three settings used recasts, prompts, and explicit correction in similar proportions, with recasts being by far the most frequent, followed by prompts then explicit correction.

The frequency of recasts in CBLT has been associated with their discourse functions that facilitate the delivery of subject matter and provide helpful scaffolding to learners when

Roy Lyster

target forms are beyond their abilities. Moreover, in the context of education, recasting has been defined more broadly than an error correction technique and rather as a scaffolding strategy that entails “the teacher’s relexicalising a student’s everyday word or words into more technical ones” (Sharpe, 2006, p. 218). In a similar vein, recasts provided in some contexts of CBLT have been described as models of more academically appropriate language (Gibbons, 2003; Mohan & Beckett, 2001).

However, there is also some evidence that recasts in other contexts of CBLT are not consistently provided for the purpose of drawing students’ attention to more academically appropriate language but rather for confirming the content or veracity of their utterances. Lyster (1998) reported that French immersion teachers repeated students’ well-formed utterances even more frequently than they recast ill-formed utterances and that, together, noncorrective repetition and recasts followed almost one-third of all student utterances, both serving to acknowledge content or to elicit additional information related to the student’s message. Teachers are known to frequently repeat students’ well-formed utterances in order to confirm referential meaning and often to “rebroadcast” the student’s message to ensure that the whole class has heard (Weiner & Goodenough, 1977).

In CBLT, noncorrective repetitions and recasts thus have the potential to converge to create contexts of pragmatic ambivalence whereby students perceive both moves as a form of positive feedback confirming the content of their message. Prompts are also susceptible to pragmatic ambivalence, as suggested by the low rates of repair following clarification requests in French immersion classrooms (Lyster & Ranta, 1997) and by Koike and Pearson’s (2005) observation that in response to clarification requests “the learners would say the same response much louder a second time” (p. 491).

Prompts, however, do not co-occur with signs of approval, whereas recasts do. Signs of approval include affirmations such as *yes*, *that’s right*, and *OK*, as well as praise markers such as *very good*, *bravo*, and *excellent*. In immersion classrooms, signs of approval were observed in equal proportions across three types of teacher responses: Approval accompanied 27% of all recasts, 26% of all noncorrective repetitions, and 29% of all teacher topic-continuation moves immediately following errors (Lyster, 1998). Thus, as documented in early studies of parent-child interaction (e.g., Penner, 1987), truth value rather than well-formedness would seem to govern approval of learner responses in CBLT.

The use of signs of approval in this way may be typical of CBLT where teachers and students alike are more focused on content than on language. In these contexts, signs of approval serve to say “yes” to content while the recasts serve to say “no” to form, with the inevitable result that some learners are more likely to notice the approval of content than the morphological modification in a recast. Moreover, Wong and Waring (2009) reported that the use of approval markers such as *very good* may inhibit learning opportunities insofar as they serve a “finale” function that precludes further attempts by others to articulate their understanding or explore alternative answers. They recommended that teachers deliver such signs of approval with a “nonfinal” tone by using “a mid-rising intonational contour, which has the effect of functioning as a continuer, soliciting ‘more’ or further responses from the students” (p. 200).

As Lyster and Ranta (1997) concluded, “Teachers might want to consider the whole range of techniques they have at their disposal rather than relying so extensively on recasts”

(p. 56). In this regard, there is some evidence that students in CBLT contexts might benefit more from feedback that pushes them to self-repair (i.e., prompts), especially in cases where recasts could be perceived ambiguously as approving their use of nontarget forms and where learners have reached a developmental plateau in their use of the nontarget forms. Notwithstanding, there is a growing consensus that students are more likely to benefit from a variety of CF types than from only one type at the expense of others (Ellis, 2012; Lyster, Saito, & Sato, 2013).

Teaching Tips

Be Aware of the Implications of Using Recasts and Prompts

- Continued recasting of what students already know is not an effective strategy for ensuring continued L2 development.
- Continued prompting of learners to draw on what they have not yet acquired will be equally ineffective.

A Proactive Approach to CBLT

The preceding section concerned a reactive approach to CBLT involving teacher scaffolding, questions, and feedback that help students to attend to language during interaction without losing sight of the content. This section addresses a proactive approach that requires planning for noticing and awareness activities followed by opportunities for guided and autonomous practice. Planning for content and language integration in this way involves shifting learners' attention to language in the context of content instruction in cases where they would not otherwise process the language at the same time as the content. This integrated approach to CBLT differs from traditional language instruction, which isolates language from any content other than the mechanical workings of the language itself.

A proactive approach to integrating language and content has been operationalized as an instructional sequence of noticing, awareness, guided practice, and autonomous practice (Lyster, 2007, 2016). The noticing activity establishes a meaningful context related to content usually by means of a text in which target features have been contrived to appear more salient (i.e., typographical enhancement) or more frequent (i.e., input flood). The awareness activity then encourages the students to reflect on and manipulate the target forms in a way that helps them to restructure their interlanguage representations. Also known in the literature as consciousness-raising tasks, awareness activities require some degree of analysis or reflection by means of rule-discovery tasks, metalinguistic exercises, and opportunities for pattern detection. The guided practice phase then provides opportunities for students to proceduralize their (re)analyzed representations of the target language in a controlled context. The sequence comes full circle at the autonomous practice phase by returning to the content area that served as the starting point. Similar to the guided practice phase, autonomous practice activities require the use of the target language features but in a disciplinary or thematic context with fewer constraints in order to encourage more autonomous use of the target language.

Key Concepts

Instructional Sequence for CBLT

Noticing: In a context related to content, students' attention is drawn to problematic L2 features highlighted through typographical enhancement.

Awareness: Students engage in some degree of metalinguistic reflection so they become more aware of the pattern.

Guided practice: Students are pushed to use the features in a meaningful yet controlled context with feedback in order to develop automaticity and accuracy.

Autonomous practice: In a context related to content, students are encouraged to use the features in more open-ended ways to develop fluency, motivation, and confidence.

To illustrate the implementation of this instructional sequence in CBLT, an example is provided here from Lyster's (2015) description of a classroom intervention with immersion students in Grade 5 (10–11 years old). Form-focused instructional activities targeting grammatical gender in French were embedded in the children's regular curriculum materials, which integrated language arts, history, and science. The research team created a student workbook that contained modified versions of texts found in the regular curriculum materials, in which noticing activities drew students' attention to noun endings as predictors of grammatical gender. For example, in the context of learning about the founding of Quebec City in 17th-century New France, endings of target nouns and their determiners had been highlighted in bold. Target words and related patterns were key to the content of the lessons. For example, *la fourrure* ("fur") was a key noun phrase because of the pivotal role of the fur trade in New France, and so was the noun phrase *la nourriture* ("food") because of the lack of food in the colony that led to a serious outbreak of scurvy.

The ensuing awareness activities required students then to detect the patterns by classifying the target nouns according to their endings and indicating whether nouns with these endings were masculine or feminine. In the case of *la fourrure* and *la nourriture*, students were expected to identify them both as feminine nouns because of their common ending *-ure*.

Then for guided practice in attributing the right gender marker to target nouns, a set of riddles was used to review the challenges experienced by settlers in New France while eliciting target nouns from students. For example, the riddle (provided in French), "*I am what covers certain mammals and can be made into warm coats,*" was intended to elicit the noun phrase *la fourrure* but, to stay in the game, a student needed to say the right gender-specific determiner, which is no small feat for young learners of French for whom grammatical gender markers, despite their frequency, are notoriously difficult.

Finally, in the autonomous practice phase, teachers returned to an emphasis on content objectives by asking students to reflect on some of the differences between life in the 17th century and life today, especially with respect to social realities and values. For example, students were asked to compare the attitudes of people in New France with those of people today concerning the fashionability of fur. Even though the subject-matter goal was to have students question and compare different social realities, teachers maintained a secondary focus on language by ensuring correct use of gender at least with key topic words such as *la fourrure*.

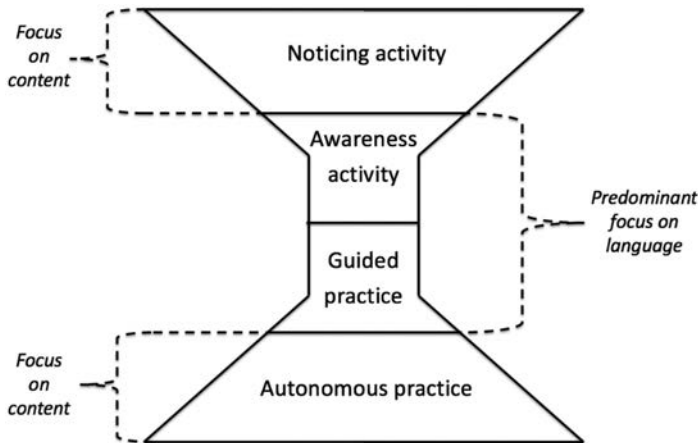


Figure 6.2 Instructional sequence integrating language and content in CBLT (adapted from Lyster, 2016, p. 58)

As illustrated in Figure 6.2, the instructional sequence begins with a primary focus on content during the noticing phase then zooms in on language during the awareness phase and guided practice phase. Finally, during the autonomous practice phase, the primary instructional focus is once again on the content that served as the starting point.

This model is reminiscent of the one proposed by Gibbons (2015, p. 227) in the shape of an hourglass to represent how mainstream teachers can focus on language as an object of study with ESL students. For teachers to do so effectively, she proposes that lessons move from whole to part, from meaning to form, and from familiar to unfamiliar.

One way for teachers in CBLT to focus more on language in some contexts than others is to distribute these activities across the language class and different content areas. Whereas the focus on language in the awareness phase and the guided practice phase might be best suited to language classes, the greater focus on content during the noticing phase and autonomous phase might be best suited to content areas. This is fairly easy for teachers to do if they teach both language and subject matter classes. If they share these responsibilities with other teachers, this is where teacher collaboration plays a key role in CBLT. For instance, in the preceding example, teachers could collaborate to plan for the noticing and autonomous practice phases to unfold in the history class, where the focus would be initially on the hardships (famine, disease, conflict) experienced by the settlers in New France, and then later on comparisons of different social realities then and now. The noticing and guided practice phases could unfold in the French class with its focus on detecting rules for grammatical gender attribution in the history texts, followed by oral practice in using target nouns with correct determiners while reviewing the history content. This type of collaboration can even serve to increase students' engagement with the content, as they perceive the involvement of two teachers rather than only one (Lyster, 2016).

A set of seven quasi-experimental studies undertaken in French immersion classrooms between 1989 and 2013 yielded overall positive effects for the integration of noticing, awareness, guided practice, and autonomous practice activities on a range of challenging target features in French (Day & Shapson, 2001; Harley, 1989, 1998; Lyster, 1994, 2004; Lyster, Quiroga, & Ballinger, 2013; Wright, 1996). In more than 75% of the 40 tests given either as immediate or delayed posttests, students participating in the form-focused tasks

Roy Lyster

improved more in their French proficiency than students left to their own devices to “pick up” the target forms from the regular curriculum (Lyster, 2016).

Teacher Collaboration Across Languages

Another type of teacher collaboration that has proven effective in contexts of CBLT is collaboration between the teachers of the two different target languages. Planning for biliteracy instruction in a way that targets both languages is based on Cummins’s (2007) argument that “learning efficiencies can be achieved if teachers explicitly draw students’ attention to similarities and differences between their languages and reinforce effective learning strategies in a coordinated way across languages” (p. 233).

To explore the feasibility of a proactive approach to cross-lingual pedagogy in the context of French immersion, Lyster, Collins, and Ballinger (2009) implemented a bilingual read-aloud project in three classrooms ranging from Grades 1 to 3 composed of French-dominant, English-dominant, and French/English bilingual students. The project aimed to facilitate collaboration between the French and English teachers of the same students as a means of reinforcing their students’ biliteracy skills. The two teachers of each class read aloud to their students from the same storybooks over 4 months, alternating the reading of one chapter from the French edition and another from the English edition. Prior to each read-aloud session, teachers asked their students to summarize the content of the previous reading, which had taken place in the other language of instruction, and after each reading they asked their students to make predictions about the next chapter, thereby generating a great deal of student interaction. Students became enthusiastic participants during the reading of the stories in both languages, which appeared to enable the students, irrespective of language dominance, to understand the stories. Many of the students continued to read stories on their own from the same book series, whether in English or French. While the read-aloud sessions led to some cross-linguistic connections made incidentally, systematic collaboration between partner teachers to make connections across languages was minimal. Based on this observation, Lyster, Quiroga, & Ballinger (2013) undertook a follow-up study designed to provide (1) more time for participating teachers to actually collaborate on planning and (2) more structured guidance regarding language objectives.

In the Lyster, Quiroga, & Ballinger (2013) study, three pairs of partner teachers (French/English) co-designed and implemented biliteracy tasks across their French and English classes at the Grade 2 level. The biliteracy tasks began in one language during its allotted class time and continued in the other language during its class time. The tasks were designed to draw attention to word formation and thereby develop students’ awareness of derivational morphology within and across languages. While the language focus was on derivational morphology, the content focus emerged from the themes of illustrated storybooks that were read in both languages.

Before and after the intervention, separate measures of morphological awareness in French and English were administered to a subsample of the students receiving the biliteracy instruction (the experimental group) as well as to a comparison group of students not receiving the instruction. At the time of posttesting, the experimental group significantly outperformed the comparison group in French, and these positive effects were similar for all students receiving the instruction irrespective of language dominance. In addition, participating teachers’ perceptions were positive and enthusiastic.

In a similar context and with similar cross-lingual objectives targeting biliteracy development, Ballinger (2013) investigated the extent to which young students can engage with a

peer in collaborative learning tasks as a means of increasing their awareness of each other's language production. The results were promising, revealing the benefits of instruction that modeled collaborative strategies (including provision of peer feedback), but that the quality of the interaction and the extent to which students' engaged in "reciprocal learning strategies" were tempered by pair dynamics (see also Ballinger, 2015).

Integrating CBLT in Foreign Language Classrooms

A counterbalanced approach to CBLT has been invoked primarily to explain the benefits of integrating a form-focused component into content-driven CBLT such as the French immersion programs in Canada (Lyster, 2007). This is because students who have been primed by their instructional setting to be meaning-oriented learners benefit from form-focused instruction designed to increase their awareness of form. Yet the converse is also true: students who have been primed by their instructional setting to be form-oriented learners benefit from content-based tasks designed to reorient their attention toward meaning. Counterbalancing their form orientation in this way is expected to contribute to their communicative abilities by averting an overemphasis on attention to form, which may jeopardize their capacity to process other equally important aspects of the input (Tomlin & Villa, 1994). In this sense, counterbalanced instruction is based on Skehan's (1998) argument that pushing learners who are either form-oriented or meaning-oriented in the opposite direction is likely to strike a balance between the two orientations in ways that promote accuracy, fluency, and complexity in target language development.

For these reasons, integrating aspects of CBLT into language-driven classrooms may prove beneficial in circumstances where the conditions for its implementation are favourable. The integration of CBLT in foreign language classrooms can be seen as a means of enriching classroom discourse for the purpose of improving language proficiency and not necessarily as a means of studying high-stakes academic content entirely through the medium of the foreign language. To illustrate counterbalanced instruction in a foreign language setting, an example is provided next of the integration of a content-based unit on environmental issues into a French as a foreign language classroom in the US (Cumming & Lyster, 2016).

A high school French teacher and her 27 students from two US foreign language classes participated in a 6-week unit on environmental issues. Data collection included measures of both language and content administered as a pretest immediately before the intervention, as an immediate posttest following the 6-week intervention, and as a delayed posttest 11 weeks later. Data collection also included classroom observations, interviews with the teacher and a subsample of students, and questionnaires administered to the teacher and all students.

The content focus on environmental issues was at the core of the instructional unit. The three-phase unit began with a focus on cause–effect patterns and then proceeded to expert–group projects in which students first researched a chosen issue in depth, using resource books made available in French, then shared their topic with other classmates by using Glogster (edu.glogster.com) to create an interactive multimedia image to teach peers about their topic. Examples of topics included fossil fuels, deforestation, climate change, air pollution, wind energy, nuclear energy, solar energy, water pollution, and overfishing. A final group project then involved the creation of a public service announcement designed to convince others to be environmentally conscious. Students' awareness of environmental issues was measured on three occasions by a task soliciting student responses in English, asking them to write a list of environmental issues they had learned about or were aware of, adding any supporting information, key terms, or other known information next to each issue.

Roy Lyster

The language focus, which was considered secondary throughout the unit, included (1) patterns of grammatical gender pertaining to key words such as *la pollution* (typical feminine ending) and *l'environnement* (typical masculine ending), and (2) informal versus formal uses of *tu/vous* imperative verb forms that students needed to use appropriately in their creation of a public service announcement. French language accuracy was assessed on three occasions through measures of grammatical gender, second person pronouns, and imperative verb forms.

The results confirmed the feasibility of integrating CBLT with foreign language instruction and yielded positive outcomes for both language and content. With respect to language, students maintained the same level of accuracy in assigning grammatical gender throughout the intervention and showed some increase in their accurate use of imperative verb forms. In terms of content, students exhibited a clear increase in their use of scientific language and level of detail. Even though the intervention unit was a content-based unit in French, students' ability to express their knowledge of environmental issues through English was enhanced.

Two other findings are also noteworthy. First, the teacher's patience in consistently using the target language in spite of the students' initial frustrations was worthwhile in the long run because it ultimately led to a motivating sense of accomplishment on the part of students. In this regard, the teacher remarked, "At the beginning, it was a little bit painful; they wanted to know the English meaning of everything, and then as we went along, they got more and more comfortable" (Cumming & Lyster, 2016, p. 88), and a student echoed her statement:

At first it was challenging, and at first you didn't get as much out of it, but as the unit went along, we learned, it was actually really beneficial to learn it in French, and to be able to understand, like, as you went along, you could just tell, everything got easier.

p. 89

Second, the content-and-language-integrated unit helped students to connect more to the language through the use of cognitively engaging and meaningful academic content. Students repeatedly mentioned how the unit applied to their own lives and was also part of a bigger picture: "It wasn't just for language—it was for science, and our world" (p. 88). These findings are in line with Wesche and Skehan's (2002) assertion that CBLT programs are "highly appreciated by students for their relevance and by teachers for the satisfaction of effectively helping students to prepare for life after language instruction" (p. 225).

Future Directions

There is a consensus in the CBLT literature that "teachers who teach content through their students' L2 require considerable professional development to effectively do so" (Lyster & Tedick, 2014, p. 219). Consequently, many of the future directions for research and development in CBLT are linked to teacher education and professional development.

The instructional integration of language and content continues to prove challenging for teachers (Cammarata & Tedick, 2012) and needs to be systematically addressed through preservice teacher education and ongoing professional development. The underlying questions include what skills teachers need in order to integrate language and content instruction effectively and also how teachers can collaborate to facilitate language and content integration. An interesting area to explore in this regard is the extent to which discipline-specific language (the language of science, mathematics, history, etc.) can be identified in ways that

help teachers integrate language and content (Llinares et al., 2012). A pivotal question that remains open for further investigation is how teachers can most effectively implement CBLT in ways that scaffold content learning while ensuring continued development in the target language.

Research on the effects of CBLT has hitherto tended to measure L2 development more than content knowledge, leaving open many questions about the feasibility and effectiveness of focusing on language during subject-matter instruction. Specifically, we still need to know whether content knowledge is enhanced or possibly compromised by a greater focus on language during content instruction. In the specific case of CLIL, it would be useful to explore closer links between the EFL class and the content class to ensure that the language addressed in the EFL class is language that complements or supports the content focus. In the specific case of two-way immersion, we need to know more about how a language focus can be adapted to accommodate different groups of learners with different language learning needs (e.g., Spanish-dominant, English-dominant, bilingual; see Tedick & Young, 2014). Finally, while there is still a need to explore effective ways of integrating a greater focus on language in content-driven classrooms, there is also a need to continue exploring ways of integrating CBLT in language-driven classrooms as a means of enriching classroom discourse and increasing opportunities for purposeful communication.

A notable strength of CBLT has been its effectiveness in the form of immersion programs supporting a variety of languages that include: (1) less widely used co-official languages (e.g., French in Canada, Swedish in Finland, Catalan in Spain, Irish in Ireland); (2) indigenous languages (e.g., Hawaiian in the US); (3) regional languages (Breton and Occitan in France); (4) heritage languages for both minority- and majority-language students (e.g., Spanish in the US); and (5) foreign or “world” languages ranging from English in Brazil and Japan to Mandarin and Japanese in the US. Along with internationalization, however, English is increasingly becoming the language targeted by many CBLT programs, most notably by CLIL (Lasagabaster & Sierra, 2010) and of course EMI (Coleman, 2006). This suggests that, for CBLT to achieve its goal of fostering rather than hindering a multilingual mindset, it needs to continue supporting languages other than only English in order to maintain the linguistic diversity that is more likely contribute to human development than convergence toward a single lingua franca (Crystal, 2000).

References

- Airey, J. (2012). I don't teach language: The linguistic attitudes of physics lectures in Sweden. *AILA Review*, 25, 64–79.
- Allen, P., Swain, M., Harley, B., & Cummins, J. (1990). Aspects of classroom treatment: Toward a more comprehensive view of second language education. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 57–81). Cambridge: Cambridge University Press.
- Arias, A., & Izquierdo, J. (2015). Language attention in content-based instruction: The case of language instructors teaching content in a foreign language in Mexican Higher Education. *Journal of Immersion and Content-Based Education*, 3(2), 194–217.
- Arnau, J., & Vila, F. X. (2013). Language-in-education policies in the Catalan language area. In J. Arnau (Ed.), *Reviving Catalan at school: Challenges and instructional approaches* (pp. 1–28). Bristol: Multilingual Matters.
- Ballinger, S. (2013). Towards a cross-linguistic pedagogy: Biliteracy and reciprocal learning strategies in French immersion. *Journal of Immersion and Content-Based Language Education*, 1(1), 131–148.

Roy Lyster

- Ballinger, S. (2015). Linking content, linking students: A cross-linguistic pedagogical intervention. In J. Cenoz & D. Gorter (Eds.), *Multilingual education: New perspectives* (pp. 35–60). Cambridge: Cambridge University Press.
- Björklund, S., Mård-Miettinen, K., & Savijärvi, M. (2013). Swedish immersion in the early years in Finland. *International Journal of Bilingual Education and Bilingualism*, 17(2), 197–214.
- Bostwick, M. (2001). English immersion in a Japanese school. In D. Christian & F. Genesee (Eds.), *Bilingual education* (pp. 125–138). Alexandria, VA: TESOL.
- Brinton, D., Snow, M., & Wesche, M. (2003). *Content-based second language instruction* (2nd ed.). Ann Arbor, MI: University of Michigan Press.
- Burger, S., & Chrétien, S. (2001). The development of oral production in content-based second language courses at the University of Ottawa. *Canadian Modern Language Review*, 58(1), 84–102.
- Cammarata, L. (Ed.). (2016). *Content-based foreign language teaching: Curriculum and pedagogy for developing advanced thinking and literacy skills*. New York: Routledge/Taylor & Francis.
- Cammarata, L., & Tedick, D. (2012). Balancing content and language in instruction: The experience of immersion teachers. *Modern Language Journal*, 96, 251–269.
- Cenoz, J. (2008). Achievements and the challenges in bilingual and multilingual education in the Basque Country. *AILA Review*, 21, 13–30.
- Coleman, J. (2006). English-medium teaching in European higher education. *Language Teaching*, 39, 1–14.
- Costa, J., & Lyster, R. (2011). Revitalization of regional languages in France through immersion. *Canadian Issues*, Fall Issue, 55–58.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge: Cambridge University Press.
- Crystal, D. (2000). *Language death*. Cambridge: Cambridge University Press.
- Cummins, J. (2007). Rethinking monolingual instructional strategies in multilingual classrooms. *Canadian Journal of Applied Linguistics*, 10, 221–241.
- Cumming, J., & Lyster, R. (2016). Integrating CBI into high school foreign language classrooms. In L. Cammarata (Ed.), *Content-based foreign language teaching: Curriculum and pedagogy for developing advanced thinking and literacy skills* (pp. 77–97). New York: Routledge/Taylor & Francis.
- Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms*. Amsterdam: John Benjamins.
- Day, E., & Shapson, S. (1996). *Studies in immersion education*. Clevedon: Multilingual Matters.
- Day, E., & Shapson, S. (2001). Integrating formal and functional approaches to language teaching in French immersion. *Language Learning*, 51(S1), 47–80.
- Early, M. (2001). Language and content in social practice: A case study. *Canadian Modern Language Review*, 58, 156–179.
- Echevarría, J., & Graves, A. (1998). *Sheltered content instruction*. Boston, MA: Allyn & Bacon.
- Echevarría, J., Vogt, M., & Short, D. (2008). *Making content comprehensible for English learners: The SIOP model*. Boston: Pearson Education.
- Ellis, R. (2012). *Language teaching research and language pedagogy*. Oxford: Wiley Blackwell.
- Feye, B., & Krzaklewska, E. (2013). *The ERASMUS phenomenon: Symbol of a new European generation?* Bern, Switzerland: Peter Lang.
- French, L. (2007). Brazilian bilingual education: Another language, another soul. *New Routes*, September Issue, 28–30.
- García Mayo, P. (2015). The interface between task-based language teaching and content-based instruction. *System*, 54, 1–3.
- Genesee, F. (2004). What do we know about bilingual education for majority language students? In T. K. Bhatia & W. Ritchie (Eds.), *Handbook of bilingualism and multiculturalism* (pp. 547–576). Malden, MA: Blackwell.
- Gibbons, P. (2015). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom* (2nd ed.). Portsmouth, NH: Heinemann.

- Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, 37, 247–273.
- Haneda, M. (2005). Functions of triadic dialogue in the classroom: Examples for L2 research. *Canadian Modern Language Review*, 62, 313–333.
- Harley, B. (1989). Functional grammar in French immersion: A classroom experiment. *Applied Linguistics*, 10, 331–359.
- Harley, B. (1998). The role of form-focused tasks in promoting child L2 acquisition. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 156–174). Cambridge: Cambridge University Press.
- Harley, B., Allen, P., Cummins, J., & Swain, M. (Eds.). (1987). *Development of bilingual proficiency: Final report: Volume II: Classroom treatment*. Toronto, ON: Modern Language Centre, OISE/UT.
- Harley, B., Cummins, J., Swain, M., & Allen, P. (1990). The nature of language proficiency. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 7–25). Cambridge: Cambridge University Press.
- Hoare, P. (2010). Content-based language teaching in China: Contextual influences on implementation. *Journal of Multilingual and Multicultural Development*, 31, 69–86.
- Hoare, P., & Kong, S. (2008). Late immersion in Hong Kong: Still stressed or making progress? In T. Fortune & D. Tedick (Eds.), *Pathways to bilingualism and multilingualism: Evolving perspectives on immersion education* (pp. 242–263). Clevedon: Multilingual Matters.
- Koike, D., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System*, 33, 481–501.
- Kong, S. (2009). Content-based instruction: What can we learn from content-trained teachers' and language-trained teachers pedagogies? *Canadian Modern Language Review*, 66, 233–268.
- Kong, S., & Hoare, P. (2011). Cognitive content engagement in content-based language teaching. *Language Teaching Research*, 15(3), 307–324.
- Lasagabaster, D., & Sierra, J. M. (2010). Immersion and CLIL in English: More differences than similarities. *ELT Journal*, 64, 367–375.
- Lindholm-Leary, K. (2001). *Dual language education*. Clevedon: Multilingual Matters.
- Lingley, D. (2006). A task-based approach to teaching a content-based Canadian studies course in an EFL context. *Asian EFL Journal*, 8, 122–139.
- Llinares, A., & Dafouz, E. (2010). Content and language integrated programmes in the Madrid Region: Overview and research findings. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain: Implementation, results and teacher training* (pp. 95–114). Newcastle upon Tyne: Cambridge Scholars.
- Llinares, A., & Lyster, R. (2014). The influence of context on patterns of corrective feedback and learner uptake: A comparison of CLIL and immersion classrooms. *Language Learning Journal*, 42, 181–194.
- Llinares, A., Morton, T., & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge: Cambridge University Press.
- Luning, R., & Yamauchi, L. (2010). The influences of indigenous heritage language education on students and families in a Hawaiian language immersion program. *Heritage Language Journal*, 7, 46–75.
- Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics*, 15, 263–287.
- Lyster, R. (1998). The ambiguity of recasts and repetition in L2 classroom discourse. *Studies in Second Language Acquisition*, 20(1), 51–81.
- Lyster, R. (2004). Differential effects of prompts and recasts in form—focused instruction. *Studies in Second Language Acquisition*, 26(3), 399–432.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.
- Lyster, R. (2015). Using form-focused tasks to integrate language across the immersion curriculum. *System*, 53, 4–13.

Roy Lyster

- Lyster, R. (2016). *Vers une approche intégrée en immersion*. Montreal: Les Éditions CEC.
- Lyster, R., Collins, L., & Ballinger, S. (2009). Linking languages through a bilingual read-aloud project. *Language Awareness*, 18, 366–383.
- Lyster, R., Quiroga, J., & Ballinger, S. (2013). The effects of biliteracy instruction on morphological awareness. *Journal of Immersion and Content-Based Language Education*, 1(2), 169–197.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37–66.
- Lyster, R., & Rebuffot, J. (2002). Acquisition des pronoms d’allocution en classe de français immersif. *Acquisition et Interaction en Langue Étrangère*, 17, 51–71.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40.
- Lyster, R., & Tedick, D. J. (2014). Research perspectives on immersion pedagogy: Looking back and looking forward. *Journal of Immersion and Content-Based Education*, 2(2), 210–224.
- Mercer, N. (1999). Classroom language. In B. Spolsky (Ed.), *Concise encyclopedia of educational linguistics* (pp. 315–319). Oxford: Pergamon.
- Mohan, B. (1986). *Language and content*. Reading, MA: Addison-Wesley.
- Mohan, B., & Beckett, G. H. (2001). A functional approach to research on content-based language learning: Recasts in causal explanations. *Canadian Modern Language Review*, 58, 133–155.
- Moriyoshi, N. (2010). *Content-based instruction in Japanese college classrooms: Focusing on language, content, or both?* (Unpublished master’s thesis). McGill University, Montreal.
- Mougeon, R., Nadasdi, T., & Rehner, K. (2010). *The sociolinguistic competence of immersion students*. Bristol: Multilingual Matters.
- Nassaji, H., & Wells, G. (2000). What’s the use of “triadic dialogue”? An investigation of teacher–student interaction. *Applied Linguistics*, 21, 376–406.
- Nicholas, H., & Lightbown, P. M. (2008). Defining child second language acquisition, defining roles for L2 instruction. In J. Philip, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the young learner: Child’s play?* (pp. 27–51). Amsterdam: John Benjamins.
- Ó Baoill, D. (2007). Origins of Irish-medium education: The dynamic core of language revitalization in Northern Ireland. *International Journal of Bilingual Education and Bilingualism*, 10, 410–427.
- Penner, S. (1987). Parental responses to grammatical and ungrammatical child utterances. *Child Development*, 58, 376–384.
- Peter, L. (2014). Language ideologies and Cherokee revitalization: Impracticality, legitimacy, and hope. *Journal of Immersion and Content-Based Language Education*, 2(1), 96–118.
- Peter, L., Hirata-Edds, T., & Montgomery-Anderson, B. (2008). Verb development by children in the Cherokee language immersion program, with implications for teaching. *International Journal of Applied Linguistics*, 18(2), 166–187.
- Poirier, J., & Lyster, R. (2014). Les pronoms objets directs de la 3e personne et leur apport aux indices du genre grammatical dans le discours oral des enseignants en immersion. *Canadian Modern Language Review*, 70(2), 246–267.
- Reedy, T. (2000). Te Reo Maori: The past 20 years and looking forward. *Oceanic Linguistics*, 39, 157–169.
- Rodgers, D. (2006). Developing content and form: Encouraging evidence from Italian content-based instruction. *Modern Language Journal*, 90, 373–386.
- Rodgers, D. (2015). Incidental language learning in foreign language content courses. *Modern Language Journal*, 99, 113–136.
- Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal*, 1, 60–73.
- Salomone, A. (1992). Immersion teachers’ pedagogical beliefs and practices: Results of a descriptive analysis. In E. Bernhardt (Ed.), *Life in language immersion classrooms* (pp. 9–44). Clevedon: Multilingual Matters.
- Sharpe, T. (2006). “Unpacking” scaffolding: Identifying discourse and multimodal strategies that support learning. *Language and Education*, 20(3), 211–231.

- Sinclair, J., & Coulthard, R. M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford University Press.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Smit, U., & Dafouz, E. (2012). Integrating content and language in higher education: An introduction to English-medium policies, conceptual issues and research practices across Europe. *AILA Review*, 25, 1–12.
- Spezzini, S. (2005). English immersion in Paraguay: Individual and sociocultural dimensions of language learning and use. *International Journal of Bilingual Education and Bilingualism*, 7, 79–98.
- Swain, M. (1988). Manipulating and complementing content teaching to maximize second language learning. *TESL Canada Journal*, 6, 68–83.
- Tan, M. (2011). Mathematics and science teachers' beliefs and practices regarding the teaching of language in content learning. *Language Teaching Research*, 15, 325–342.
- Tedick, D., & Young, A. (2014). Fifth-grade two-way immersion students' responses to form-focused instruction. *Applied Linguistics*, 1–25. doi:10.1093/applin/amu066
- Tedick, D., & Wesely, P. (2015). A review of research on content-based foreign/second language education in US K–12 contexts. *Language, Culture and Curriculum*, 28(1), 25–40.
- Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183–203.
- Trent, J. (2010). Teacher identity construction across the curriculum: Promoting cross-curriculum collaboration in English-medium schools. *Asia Pacific Journal of Education*, 30(2), 167–183.
- Weiner, S. L., & Goodenough, D. R. (1977). A move toward a psychology of conversation. In R. Freedle (Ed.), *Discourse production and comprehension* (pp. 213–225). Norwood, NJ: Ablex.
- Wesche, M. (Ed.). (2001). French immersion and content-based language teaching in Canada [Special Issue]. *Canadian Modern Language Review*, 58(1).
- Wesche, M., & Skehan, P. (2002). Communicative, task-based, and content-based language instruction. In R. Kaplan (Ed.), *The Oxford handbook of applied linguistics* (pp. 207–228). Oxford: Oxford University Press.
- Wong, J., & Waring, H. Z. (2009). “Very good” as a teacher response. *ELT Journal*, 63(3), 195–203.
- Wright, R. (1996). A study of the acquisition of verbs of motion by grade 4/5 early French immersion students. *Canadian Modern Language Review*, 53, 257–280.

Task-Based Language Teaching

Rod Ellis

Background

The concept of ‘task’ is central to an understanding of task-based language teaching (TBLT); consequently, this section begins with a definition of a task. There follows a brief exposition of how TBLT developed out of communicative language teaching (CLT), followed by a comparison of task-supported language teaching (TSLT) and TBLT. This opening section concludes with a statement of the aims of the TBLT.

Defining ‘Task’

To understand what is meant by a task, it is important to distinguish the *task-as-workplan* from the *task-as-process* (Breen, 1989). The former consists of the instructional materials that make up the task—typically some kind of verbal or nonverbal input and a rubric that specifies what outcome the learners are asked to achieve. For example, the Heart Transplant Task provides learners with information about four people in need of a heart transplant and asks the learners to decide which of the four is most deserving of a transplant if only one heart is available. The task-as-process is the activity that transpires when learners perform the task. It involves learners exchanging information about the four people, evaluating the merits of each one, reaching a decision about who should get the transplant, and giving their reasons. Seedhouse (2005) pointed out that it is not possible to make precise predictions about what processes result from a workplan although, as we will see, the design of the workplan can lead to identifiable effects on how the task is performed.

Key Concepts

Task-as-workplan: The task teaching materials.

Task-as-process: The actual performance of the task.

Ellis and Shintani (2014) suggest that for a workplan to qualify as a task it must satisfy four criteria:

1. The primary focus should be on ‘meaning’ (i.e., learners should be concerned mainly with encoding and decoding messages, not with focusing on linguistic form).
2. There should be some kind of gap (i.e., a need to convey information, to express an opinion, or to infer meaning).
3. Learners should rely largely on their own resources (linguistic and nonlinguistic) in order to complete the task. That is, learners are not taught the language they will need to perform a task although they may be able to borrow from the input the task provides to help them perform it.
4. There is a clearly defined outcome other than the use of language for its own sake. Thus, when performing a task, learners are not primarily concerned with using language correctly but with achieving the goal stipulated by the task.

Using these criteria it is possible to distinguish between a ‘task’ and an ‘exercise.’ The Heart Transplant Task is clearly a task because it satisfies all four criteria. In contrast, instructional materials of the blank-filling kind clearly do not satisfy the requirements because the focus is primarily on form, there is no gap, learners need to draw on their resources only to insert a word or two into a ready-made sentence, and there is no outcome other than the completed exercise. However, as Ellis (2010) pointed out, some instructional materials may satisfy some but not all of these criteria and thus constitute workplans that lie on the continuum between a ‘task’ and an ‘exercise.’ Also, there are other definitions of a task that differ somewhat from those mentioned. Willis and Willis (2007), for example, proposed that a task should relate to real-world activities (i.e., match the real-life tasks that people perform and thus manifest situational authenticity).

Origins of Task-Based Language Teaching

Task-based language teaching (TBLT) is a development of communicative language teaching (CLT), which emerged in the late 1970s as an alternative to more traditional structure-based approaches to language teaching. Johnson (1982), for example, advocated what he called the ‘deep end strategy,’ where the student is asked to perform a communicative task even though he or she may need to use language that has not yet been taught. In early CLT, however, communicative tasks were seen as a means of developing fluency and, as such, were viewed as adjuncts rather than alternatives to accuracy-oriented activities such as fill-in-the-gap and substitution exercise.

Subsequently, CLT evolved into a weak and a strong form (Howatt, 1984). The weak form, like earlier approaches, was based on an inventory of the structural properties of the target language and a methodology consisting of presentation-practice-production (PPP), with tasks serving as the means for the production stage. The emphasis is on learning-to-communicate. In contrast, the strong form was based on a syllabus consisting of tasks and a methodology that emphasized learning-through-communication. In other words, the weak form of CLT entailed *task-supported language teaching* and the strong form *task-based language teaching*. Both approaches involved tasks, but the tasks functioned in a fundamentally different way.

Key Concepts

Communicative Language Teaching

- **Weak form:** Use of tasks in a structural approach to teaching (i.e., task-supported teaching).
- **Strong form:** Tasks serve as the basis for the teaching syllabus (i.e., task-based teaching).

Task-Supported and Task-Based Language Teaching Compared

Table 7.1 provides a more detailed specification of the differences between task-supported and task-based language teaching. A fundamental difference lies in how they handle attention to form. In TSLT the learners' attention is directed to the specific target form that is the focus of a lesson in the presentation stage of PPP, often by means of explicit description. In TBLT attention to form occurs while learners are performing a task, either when a communication problem occurs that leads to attention being paid to form or because one of the task participants (for example, the teacher) chooses to draw attention to a linguistic form. In TBLT, however, attention to form is always secondary to the primary aim of communicating in order to achieve the outcome of the task. Thus, whereas in TSLT the learner's primary focus is on accurate use of the target form, in TBLT the primary focus is on the communicative use of language, and attention to form is secondary.

TSLT and TBLT also cater to different kinds of learning. In TSLT learners are made aware of what linguistic forms they are supposed to learn and so the learning that takes place is intentional (i.e., learners are expected to try to understand the target feature and to use it correctly). In contrast, TBLT caters to incidental learning (i.e., the picking up of words and structures while the learner's attention is focused primarily on meaning). In this respect, TBLT aims to replicate the natural learning that takes place during first language acquisition. However, when learners learn incidentally, they may still pay conscious attention to linguistic form while they are communicating. In other words, incidental learning is not the same as implicit learning, where learning takes place without conscious awareness of linguistic form.

The theoretical basis of TSLT is skill-learning theory, which claims that learning commences with a declarative representation of a skill (or in the case of language, a linguistic

Table 7.1 Comparison of task-supported and task-based language teaching

	<i>Task-supported language teaching</i>	<i>Task-based language teaching</i>
Syllabus	Structural (i.e., a graded list of linguistic features to be taught)	Task-based (i.e., a graded list of tasks or task-types to be performed)
Attention to form	Directs attention to form	Attracts attention to form
Activity type	Exercises + tasks	Tasks only
Primary focus	Accurate use of target forms	Communicative use of language
Type of learning	Intentional	Incidental
Theory of language learning	Skill-learning theory	Interaction approach; usage-based learning
Educational philosophy	Transmission: learning-to-do	Experiential: learning-by-doing

form) that transforms through practice into procedural knowledge. This process is accompanied by a switch from controlled to automatic processing. DeKeyser (1998) emphasized that learners need the opportunity to engage with the target feature under *real operating conditions* to achieve automatic processing. This is why tasks play an essential role in TSLT; they create the conditions for using the target feature in a communicative task after declarative knowledge of the target feature has been established by means of explicit instruction. TBLT, on the other hand, draws on a variety of theories in second language acquisition research, in particular the Interaction Approach (Gass & Mackey, 2007) and usage-based language learning (N. Ellis, 2005). According to the Interaction Approach, which draws on Long's (1996) *Interaction Hypothesis*, interaction that helps to make input comprehensible, provides feedback on learners' attempts to use the language, and pushes learners to modify their own output to make it more target like assists 'natural' learning and helps learners to acquire the kind of linguistic knowledge they need to engage in communication. Usage-based theories view language learning as a gradual process that starts with ready-made chunks of language that are disassembled and combined to construct more abstract constructions that are rule-like but not exactly rule-based. Reflecting this position, in TBLT there is no attempt to teach learners declarative knowledge of target features prior to the performance of a task. These theories offer fundamentally different views of how an L2 is learned, as reflected in the fundamental differences between TSLT and TBLT.

Key Concepts

Real operating conditions: The conditions that prevail when language is used for natural communication (e.g., involve automatic processing).

Interaction Hypothesis: Claims that the negotiation of meaning that occurs when learners experience a communication problem facilitates acquisition (Long, 1996).

Finally, TSLT and TBLT can be distinguished in terms of the educational philosophies that underpin them. To a large extent TSLT draws on traditional views of classroom learning that emphasize transmission of the established facts about language. Language is dissected into discrete elements that can then be taught bit by bit. In effect, learners must first learn these bits before they can use them. TBLT accords with Dewey's (1938) emphasis on active discovery through problem solving and the importance of a holistic, learner-driven pedagogy. It involves learning through doing by creating "experience-based opportunities for language learning" (Samuda & Bygate, 2008, p. 36). Long (2015) aligns TBLT with philosophies of education that emphasize guided individual freedom to learn, emancipation, learner-centredness, egalitarian teacher–student relationships, participatory democracy and the natural human inclination to behave cooperatively.

Aims of Task-Based Language Teaching

TBLT, then, is an approach to teaching a second/foreign language that seeks to engage learners in natural language use and promote acquisition by having them perform a series of communicative tasks. In TBLT learners are encouraged to treat language as a tool for making meaning rather than as an object to be studied, practiced and learned. TBLT aims to create contexts where learners can utilize their existing linguistic resources in communication and

Rod Ellis

in this way develop fluency in the use of the L2. At the same time, TBLT also aims to help learners acquire new linguistic knowledge incidentally from both the input and interactions that tasks create, as well as the attention to form that arises naturally from the performance of a task. In other words, through performing tasks learners develop both linguistic and interactional competence in an L2. TBLT is a type of teaching that emphasizes learning through experiencing the use of the L2.

Current Issues

As described earlier, the rationale for TBLT draws on both psycholinguistic accounts of how an L2 is learned and general educational principles. Its strength lies in what Long (2015) called the ‘synergistic relationship’ between these two bodies of thinking. It is therefore, not surprising that TBLT has attracted increasing attention over the last two decades from both language educators such as Willis (1996) and SLA researchers such as Long (1985, 2015) and Skehan (1998) to the point where TBLT has achieved the status of an established approach—recognized as such by its inclusion in the most recent edition (2014) of Richards and Rogers’s *Approaches and Methods in Language Teaching*—and worthy of its own series of research-oriented books published by John Benjamins.

The danger is that TBLT is now understood to comprise an agreed set of principles and procedures that its advocates all adhere to. However, this is far from the case. There are in fact different versions of TBLT as shown in Table 7.2, which shows how different advocates of TBLT position themselves with regard to a number of key features of TBLT. These key features follow:

Table 7.2 Differences in task-based language teaching approaches

Features	Long (1985, 2015)	Willis (1996) and Willis and Willis (2007)	Skehan (1998)	Ellis (2003)
Natural language use	Yes	Yes	Yes	Yes
Course design	Target-tasks → pedagogic tasks	Pedagogic tasks	Pedagogic tasks	Pedagogic tasks
Task type	Primarily <i>unfocused tasks</i> (i.e., tasks not aimed at eliciting specific target features)	Unfocused	Unfocused	Unfocused and <i>focused tasks</i>
Task modality	Output-based	Output-based	Output-based	Both input-based and output-based
Focus on form	Yes—main task phase (negotiation of meaning)	Yes—posttask phase	Yes—pretask phase (strategic planning)	Yes—all phases
Learner-centeredness	Yes	Yes	Yes	Not necessarily
Rejection of traditional approaches	Yes	Yes	Yes	No

1. What is common to all four approaches is the emphasis on natural language use. That is, TBLT aims to promote language learning by means of tasks that create interactionally authentic contexts for the use of language. In other respects, however, the four approaches differ.
2. Long (1985, 2015) argues that the design of a task-based course should start from a needs analysis to identify the *target tasks* that a specific group of learners will need to master. *Pedagogic tasks* are then developed from the target tasks. In contrast, Ellis (2003), Skehan (1998), and Willis (1996) see no need to take target tasks as the starting point and instead propose that a course be composed of pedagogic tasks matched to learners' developing language proficiency.
3. Tasks can be unfocused (i.e., designed to elicit general samples of language use) or focused (i.e., designed to provide a communicative context for the use of specific linguistic features, such as a set of words or a particular grammatical feature). Only Ellis (2003) suggests that some tasks can be of the focused kind.
4. In general advocates of TBLT view tasks as creating opportunity for language production (i.e., as output based). Ellis, however, has argued that input-based tasks (i.e., tasks involving listening or reading) have an important role to play in TBLT, especially for beginner level learners.
5. All four approaches recognize that a *focus-on-form* is a necessary feature of TBLT, but they differ in how this should be achieved. Long sees a focus on form arising primarily out of the negotiation of meaning that takes place when a communication problem arises. Willis relegates attention to form to the posttask phase of a lesson and insists that in the main task phase (i.e., when the task is being performed) the focus should be entirely on meaning. Skehan emphasizes the importance of planning in the pretask phase of the lesson as a way of enabling learners to pay greater attention to form when they perform the task. Ellis sees opportunities for a focus on form in all phases of a task-based lesson.
6. TBLT is generally characterized as a learner-centred approach with learners performing tasks interactively in small groups. This is reflected in Long's, Willis's, and Skehan's accounts of TBLT. Ellis, however, does not see group work as an essential feature of TBLT, arguing that tasks can be performed in a whole-class context with the teacher functioning as a participant in the task.
7. Advocates of TBLT tend to dismiss traditional approaches to language teaching such as PPP. Ellis, however, suggests that a modular approach is possible, with TBLT and traditional, language-centered approaches constituting separate and unconnected modules in a complete course. In this respect he is more in line with the role assigned to tasks in early CLT (see earlier in this chapter).

There is perhaps enough commonality in the four approaches to justify the claim that they are all task-based (as opposed to task-supported), but any discussion of the issues involved in TBLT must take account of the differences.

Key Concepts

Target tasks: The tasks that people perform in real life and which according to Long serve as the basis for the design of a task-based course.

Pedagogic tasks: Pedagogic workplans that may or may not be based on target tasks.

Rod Ellis

Not surprisingly because TBLT constitutes a radical departure from traditional approaches to language teaching based on a linguistic syllabus, it has aroused considerable criticism. However, many of the critiques derive from misconceptions about TBLT and, in particular, from the failure to recognize that TBLT does not prescribe a narrow set of techniques and does not constitute a totally uniform way of teaching—as shown in Table 7.2. In Table 7.3 I have listed some of the main criticisms and my responses to them (see also Ellis, 2009a).

Table 7.3 Addressing some common misconceptions about TBLT

Criticism	Response
Seedhouse (2005) argued that ‘task-as-workplan’ has weak construct validity because the interaction that transpires when learners perform a task (i.e., the ‘task-as-process’) frequently does not match that intended by designers of the task.	While it is not possible to predict with precision what activity will result from the performance of a task-as-workplan, there is now sufficient evidence to show that the design of tasks and the way they are implemented can lead to predictable effects on performance (see next section of this chapter).
Tasks are seen primarily as a means for developing communicative fluency (Klippel, 1984).	This constitutes a fundamental misconception of the role of tasks. TBLT aims to develop both communicative fluency and linguistic accuracy. Tasks foster the incidental acquisition of new vocabulary and grammar.
Sheen (2003) argued that TBLT requires that any treatment of grammar take the form of quick corrective feedback allowing for minimal interruption of the task activity.	A <i>focus-on-form</i> (including grammatical form) can be achieved in a number of ways in TBLT: <ul style="list-style-type: none"> • Through pretask activities such as strategic planning. • Through corrective feedback during the performance of a task. • Through posttask activities, which can include direct teaching of any language items the learners experienced problems with while performing the task.
Littlewood (2007, p. 244) pointed out that speaking tasks are difficult for learners of low proficiency and may result in ‘minimal demands on linguistic competence’ and thus TBLT is not suited for beginner level learners.	Tasks can be input-based or output-based and can involve all four skills. For beginner-level learners input-based tasks that do not require production are most appropriate and can establish a basis for later production tasks.
TBLT requires extensive use of group-work, which may not be appropriate in some teaching contexts. Carless (2004), for example, reported that primary school teachers in Hong Kong experienced difficulty with group-based TBLT because students relied on their L1 and made too much noise.	Although group work is important it is not an essential feature of TBLT. Input-based tasks will need to be carried out in a whole-class context while information-gap tasks can also be performed in this way. The Communicational Language Teaching Project (Prabhu, 1987), the first attempt to implement a TBLT course, did not involve any group work.
Swan (2005) claimed that “the thrust of TBLT is to cast the teacher in the role of manager and facilitator of communicative activity rather than an important source of new language” (p. 391).	Again, this assumes that TBLT involves learners performing tasks in groups. In fact, TBLT requires the teacher to perform a variety of roles including those of manager and facilitator of communication but also the traditional roles of corrector and provider of new language.

Criticism	Response
In TBLT the teacher and the students should avoid use of the learners' first language (L1) in order to maximize exposure to the L2 (Prabhu, 1987).	The L1 has a role to play in TBLT. Learners have been shown to make effective use of the L1 to establish the goals for a task and the procedures to be followed in tackling it. Learners have also been seen to use the L1 to scaffold production in the L2 (Anton & DeCamilla, 1998).
Swan (2005) argued that TBLT is suited only to 'acquisition-rich' environments (i.e., second language contexts) and not to 'acquisition-poor' environments (i.e., foreign language contexts), where a more structured approach is required to ensure the grammatical resources needed for communicating.	Arguably TBLT is more suited to the foreign language (FL) classroom than the second language classroom as FL learners have few opportunities to communicate outside the classroom so need them inside. In contrast, learners in second language contexts have opportunities to communicate outside the classroom, so, arguably, instruction could focus more on linguistic accuracy.
Swan (2005) claimed that "in the tiny corpus of a year's task-based input, even some basic structures may not occur often, much core vocabulary is likely to be absent, and many other lexical items will appear only once or twice" (p. 392). In this respect TBLT is inferior to traditional structure-based approaches.	This misconception appears to derive from the view that tasks must inevitably involve spoken interaction and oral production. But, in fact, tasks can also be input-based (i.e., involve listening or reading). Indeed, extensive reading activities can serve as a basis for tasks. Arguably, a task-based course is capable of providing much greater exposure to the target language than a traditional course.

Key Concept

- **Focus-on-form:** The attention to form that learners pay while primarily engaged in the effort to communicate meaningfully.

There are, however, a number of issues that are more problematic. Widdowson (2003) claimed that "the criteria that are proposed as defining features of tasks are . . . so loosely formulated . . . that they do not distinguish tasks from other more traditional classroom activities" (p. 126). While this criticism is unwarranted, as the four criteria proposed earlier can distinguish a task and an exercise (see Ellis, 2010), it would appear that teachers do experience difficulty in determining whether an instructional activity is a task. Carless (2004), for example, reported that the primary school teachers he investigated in Hong Kong did not always have a clear understanding of what a 'task' was, and as a result their tasks ended up as 'language practice' rather than affording opportunities for genuine communication. Erlam (2016), too, reporting on a course for experienced teachers of foreign languages in New Zealand, found that only 20 out of the 43 tasks that the teachers developed fulfilled all four criteria. She also noted that the most difficult criterion to satisfy was the third—the need for learners to rely on their own resources (instead of being provided with the language needed to perform the task), with only 27 of the tasks meeting this criterion. However, Erlam reported that 87% of the tasks satisfied at least three of the criteria. A task-like activity may suffice to ensure that it results in 'natural language use.'

Other significant issues concern the problems that teachers and students may face in implementing TBLT in particular teaching contexts. Littlewood (2007) argued that any approach must take account of the cultural context in which teaching takes place and the particular teachers and learners involved. An experiential approach is likely to face resistance from teachers and learners who are accustomed to a transmission-based approach. TBLT threatens the established role of teachers by repositioning them as co-communicators rather than as sources of knowledge. Teachers who lack confidence in their own L2 proficiency may be especially reluctant to take on TBLT. Also, TBLT is unlikely to find much support in a context where the high-stakes language tests encourage discrete-point teaching and memorization. It was for reasons such as these that Littlewood advocated task-supported teaching. An alternative approach—the one advocated by Ellis (2003)—is a *modular curriculum*, with one module of the curriculum consisting of TBLT and another, completely separate module based on more traditional approaches. Such a curriculum acknowledges the attested value of formal, explicit instruction but also provides for the development of the interactional competence learners will need to communicate effectively in the real world. An additional advantage is that the inclusion of more formal types of instruction will also pose less of a threat to teachers used to traditional approaches.

Key Concept

Modular curriculum: A curriculum that consists of separate and unrelated components (e.g., a structural and a task-based component).

Another very real concern is how to construct a task-based course. This involves deciding which tasks (and which type of task) to include and, crucially, how the tasks can be effectively sequenced to ensure a progression from ‘easy’ to ‘difficult.’ The next section will examine how some researchers have attempted to tackle this issue.

Both Sheen (2003) and Swan (2005) have argued that there is no research to show that TBLT is more effective than traditional approaches, and, therefore, advocates of TBLT are guilty of ‘legislating by hypothesis’ (i.e., the critical application of untested theories of L2 acquisition to language pedagogy). This criticism, however, is also unfounded. To demonstrate why it is unfounded, it is necessary to consider the empirical evidence in support of TBLT.

Empirical Evidence

It is helpful to distinguish three types of empirical evidence that lend support to TBLT. The first kind involves research that made use of tasks to investigate hypotheses drawn from SLA theories. This type of evidence does not speak directly to TBLT but rather provides empirical support for the general principles that underpin TBLT. As such it addresses Swan’s (2005) complaint that TBLT is founded on hypotheses for which there is no empirical evidence. The second kind of research is more directly concerned with TBLT but is limited to investigating the performance resulting from different types of tasks. This research is important as it provides information that can be used to select and grade tasks and to develop a methodology for implementing tasks. The final type of research consists of experimental comparisons of the learning processes and the learning outcomes of TBLT and traditional

types of instruction (such as PPP). The following sections provide an introduction to these three types of research.

Tasks in SLA Research

Tasks have served as one of the main ways of researching L2 acquisition. *Focused tasks*, for example, have been widely employed in form-focused instruction studies to see whether instruction directed at a specific target feature has any effect on learners' ability to use that feature spontaneously in communication. It is not an overstatement to say that much of what we currently know about L2 acquisition has been obtained through the analysis of data collected by means of tasks of various kinds.

Key Concepts

Unfocused tasks: Tasks designed to elicit processing of general samples of language.

Focused tasks: Tasks designed to elicit the processing of some predetermined linguistic feature (e.g., a specific grammatical structure).

The research to date does support the three hypotheses that Swan (2005) saw as underpinning TBLT. There is plenty of evidence, for example, to support the 'online hypothesis,' which claims that incidental learning can take place online while learners are performing tasks. Mackey and Goo (2007) reported a meta-analysis of 28 sample studies where tasks were used to generate interaction involving L2 learners. Immediate posttests showed that the overall effect of the opportunity to interact (compared with no such opportunity) was large and also increased over time (i.e., in delayed posttests). Similarly, there are studies that lend support to the *Noticing Hypothesis*, which claims that conscious attention to form is needed for learning to take place. Learners do notice linguistic forms in the input or in the feedback they receive on their own attempts to use the L2 and such noticing is related to learning (e.g., Loewen, 2005; Mackey, Gass, & McDonough, 2000). The third hypothesis, the *Teachability Hypothesis* (Pienemann, 1985), states that the direct teaching of a grammatical structure will result in acquisition only if the learner is developmentally ready to acquire it. Traditional teaching based on a structural syllabus is incompatible with this hypothesis given the practical problems of determining whether the learners in a particular class are ready to acquire the target structure of the lesson. However, it is compatible with TBLT, as SLA research has shown that acquisition is a gradual and dynamic process and there are constraints on what learners are able to acquire at particular times. TBLT acknowledges this by seeking only to attract rather than direct attention to form and in this way caters to the natural, organic way in which learners' L2 systems develop. Swan (2005) argued that there is a lack of 'wide-ranging empirical evidence' (p. 381) to support the Teachability Hypothesis but in fact there is clear evidence that learner-readiness determines whether explicit instruction is effective (see Ellis, 2002, 2015 for a review of this research). Explicit instruction is sometimes successful but researchers are not yet in a position to specify the conditions that determine whether or not it will work. Thus the 'teachability' of a new target feature remains a problem for traditional instructional approaches including task-supported language teaching.

Key Concept

Noticing Hypothesis: This states that for learning to take place learners need to pay conscious attention to concrete linguistic elements in the input they are exposed to.

Researching Tasks

In SLA tasks or task-like activities (i.e., activities manifesting some but not all the defining characteristics of a task listed earlier in this chapter) served initially as elicitation devices for investigating L2 acquisition, but starting in the 1980s they became an object of enquiry in their own right often with pedagogy in mind. Early studies of tasks (e.g., Tong-Fredericks, 1984) were motivated simply by a desire to find out what kind of language use resulted from different tasks. Later studies were more focused on how specific design features affected the nature of the interactions that took place when learners performed tasks with native-speakers or with other learners (see Pica, Kanagy, & Falodun, 1993). Other studies (e.g., Foster & Skehan, 1996) investigated how different design features (e.g., whether the information in the task was loosely or tightly structured) and different implementation options (e.g., whether or not learners had the opportunity to plan before they performed a task) impacted on the complexity, accuracy, and fluency of the language that learners produced. Pedagogically minded researchers such as Pica and Skehan felt that such research was necessary to provide a principled basis for the development of task-based language programmes.

Task-based research has blossomed since. To make sense of this research it is helpful to first consider three sets of variables: (1) task design variables, (2) task implementation variables, and (3) aspects of the language use resulting from the performance of tasks. The general goal of this research was to try to establish what effect different design and implementation variables had on the different aspects of language use when a task was performed.

In the first instance researchers were interested in the effect that design and implementation variables had on the negotiation of meaning, but much of the later research investigated their effects on three aspects of language use: complexity, accuracy, and fluency (CAF). Complexity refers to the extent to which learners produce complex constructions and is considered to demonstrate that they are taking the risks that will lead to ‘restructuring’ of their L2 systems. Accuracy concerns the extent to which learners conform to target language norms and avoid making errors. Fluency is the extent to which learners can speak rapidly without undue pausing, repetition, or reformulation. Various ways of measuring these three constructs have been developed (see, for example, Ellis & Barkhuizen, 2005; Housen, Kuiken, & Vedder, 2012).

Table 7.4 lists the main task design variables that have figured in the research to date. All of these variables have been found to impact on CAF in ways that are to some extent predictable. For example, tasks with split information tend to result in a higher frequency of negotiation sequences than tasks with shared information. Tasks with familiar content promote fluency and accuracy. Dialogic tasks encourage greater accuracy and complexity but lower fluency. Tasks with many elements to be manipulated lead to more complex language. Reviews of the relevant research that have investigated these design variables can be found in Ellis (2003), Robinson (2011), and Skehan (2001).

Table 7.4 Typical design variables

<i>Design variables</i>	<i>Commentary</i>
1. Dialogic vs. monologic	A dialogic task requires two or more participants to interact when performing the task. A monologic task requires the individual learner to perform the task without interruption.
2. Number of elements to be manipulated	The task may require only a few elements to be communicated (e.g., in a story with just two characters in one setting) or many elements (e.g., in a story involving a number of characters in different settings).
3. Topic familiarity	A familiar topic is one where the participants have a ready-made schema they can draw on (e.g., describing the route they follow from school to their home).
4. Shared vs. split information	In a shared information task all the participants have access to the same information; alternatively the information to be communicated can be split between the participants. The former occurs in opinion-gap tasks and the latter in information-gap tasks.
5. Single vs. dual task	The difference here concerns whether the task poses learners a single goal (e.g., to draw a route on a map) or a dual goal (e.g., to draw a route on a map when the map does not correspond exactly to the route being described).
6. Closed vs. open outcome	Tasks with a closed outcome have a single solution (e.g., the route drawn on a map). Tasks with an open outcome have several possible solutions. Information-gap tasks typically have closed outcomes whereas opinion-gap tasks have open outcomes.
7. Discourse mode	The task may lead to discourse involving description, instructions, narrative, or argument.
8. Here-and-now vs. there-and-then orientation	Tasks may require participants to refer to entities and actions that they can see occurring (as when they describe a live video) or to entities and actions that are not physically present (as when they describe a video they have just watched).

Task planning is the implementation variable that has received the greatest attention (see Ellis, 2005; Skehan, 2014). Here it is useful to distinguish two types of planning—pretask planning and online planning. The former has been investigated by comparing the task performances of learners who are given time to plan what they want to say before they undertake the task with learners who start straight in with the task without any planning time. Pretask planning, as might be expected, enhances fluency but it also tends to benefit either complexity or accuracy (but not both). Online planning is operationalized in terms of whether or not learners are asked to perform the task under time pressure. Research suggests that learners' accuracy increases when they have ample time for online planning. Ellis (2009b) reviewed the research that has investigated these two types of planning along with task-repetition (i.e., asking learners to repeat the same task).

Teaching Tips

- To encourage learners to use complex language and also to speak fluently, allow time for them to plan before they start performing the task.
- To encourage learners to focus more on accuracy, allow them to perform the task without any time pressure.

To a large extent the early task-based research was exploratory in nature. The later research, however, was theoretically driven, in particular by Skehan's (1998) Limited Attention Capacity Hypothesis and Robinson's (2001) Cognition Hypothesis. These hypotheses draw on different models of working memory. Skehan argued that limitations in learners' working memory makes it difficult for them to pay attention simultaneously to both meaning and form and that they will therefore tend to prioritize one or the other depending on the task conditions. In particular Skehan suggested that there is likely to be a trade-off between complexity and accuracy (i.e., learners will find it difficult to produce language that is both complex and accurate and thus they will prioritize one depending on the design of the task and how it is implemented). Robinson's Cognition Hypothesis is more ambitious, aiming to account for how task complexity, interactive conditions, and individual learner factors impact on task performance. Robinson argued that working memory is expandable and that more complex tasks will result in language that is both more complex and more accurate (i.e., there is no trade-off). There are studies that lend support to both theories, but in general the task-based studies to date lend greater support to Skehan's position. Jackson and Seuthanpronkul (2013), for example, failed to find support for a dual effect on language complexity and accuracy in a meta-analysis of nine task-based studies.

The task-based research has been ambitious in trying to show how specific design and implementation variables have particular effects on interaction and CAF. These studies have undoubtedly provided some important insights. However, this approach is not without its problems. As Skehan (2014) pointed out "any task is likely to subsume a bundle of features" (p. 6). One can ask therefore whether it is really possible to isolate the effect of specific task variables as so much of the research has attempted to do. Questions have also been raised about the types of CAF measures used to investigate task performance (see Lambert & Kormos, 2014) and the failure to obtain independent measures of task complexity, for example by asking learners about their perceptions of a task after they have performed it (see Revesz, 2014).

Comparative Studies of PPP and TBLT

To address the relative effectiveness of traditional approaches and TBLT, comparative studies of the two types of instruction are needed. There have in fact, been few such studies. The best comparative study to date is Shintani (2011, 2015), which compared the relative effects of TBLT consisting of input-based tasks and presentation-practice-production (PPP) on the acquisition of a set of new words and the incidental acquisition of two grammatical structures (plural *s* and copula *be*) by young beginner Japanese learners of L2 English in Japan. The results indicated that both types of instruction were effective but that overall TBLT was superior. The TBLT learners performed just as well on tests of the words targeted in the PPP

lessons and included in the input-based tasks; however, they also learned more of the words that arose incidentally in the two types of instruction. Furthermore, only the TBLT learners acquired plural-*s* incidentally. Shintani also showed that the interactions that occurred in the two types of instruction were fundamentally different. The input-based lessons lead to opportunities for learners to initiate discourse and to negotiate for both meaning and form. In contrast, the PPP lessons resulted primarily in initiate-respond-feedback (IRF) exchanges that are so ubiquitous in formal instruction.

Shintani's study lends support, then, to the fundamental claim of TBLT—namely that TBLT facilitates the concomitant development of linguistic and interactional competence and that it can achieve this more effectively than PPP. Clearly, however, more well-designed comparative studies are needed, especially of older learners who might be expected to benefit more fully from traditional approaches.

Pedagogical Implications

While most of the task-based research has focused on the performance of individual tasks, task-based pedagogy needs to take a broader perspective by considering the design of complete task-based courses and the organization of task-based lessons. There is, therefore, a gap between the main focus of the research and pedagogy.

There are, however, a number of research-based proposals for designing task-based courses. In Prabhu's (1987) Communicational Language Teaching Project in India, information-gap tasks worked best for beginner-level learners with reasoning tasks and opinion-gap tasks were better suited to more proficient learners. Ellis (2003) pointed out that a natural sequence for a task-based course would be to start with input-based tasks and then move on to output-based tasks. Skehan (1998) suggested that a curriculum could achieve a balance in the development of complexity, accuracy, and fluency by designing tasks that biased learners toward these different aspects of language use. The chapters in Van den Branden (2006) also offer practical suggestions for how research-based principles can be applied in the development of task-based teaching materials.

These suggestions are helpful, but there remains the key issue about how to sequence tasks in a task-based course. The most detailed proposal for addressing this has come from Robinson and his co-researchers (see for example, Baralt, Gilabert, & Robinson, 2014). This proposal draws on Robinson's Cognition Hypothesis, which supports an ordering of tasks in terms of their cognitive complexity determined by a set of 'resource-directing factors' (such as those in Table 7.4), which govern the extent to which learners attend to form while performing a task. In contrast, resource-dispersing variables such as pretask planning simplify a task and thus promote fluency. Robinson advanced his SSARC Model (simplify, stabilize, automatize, restructure, complexify) for sequencing different versions of the same task. The starting point is to simplify a task in terms of both resource-directing and resource-dispersing variables (e.g., +pretask planning; −reasoning), then increase complexity first in terms of resource-dispersing variables (e.g., −pretask planning; −reasoning) and finally in terms of both resource-dispersing and resource-directing variables (e.g., −pretask planning; +reasoning). Robinson argues that sequences of simple to complex tasks help to remind learners of previous learning episodes and thereby consolidate memory for them. The problem with Robinson's proposal is the same as with the Cognition Hypothesis—there is insufficient evidence to show that task complexity (established in accordance with the resource-directing variables) has the joint effect on complexity and accuracy of language use that Robinson predicts. To date, there is no published course textbook based on Robinson's SSARC Model.

Rod Ellis

It is difficult to see the research providing a scientific basis for sequencing tasks. Certainly, we are a long way from this as things currently stand. Tasks are holistic in nature and involve clusters of features. Thus a task that is ‘complex’ in terms of one feature (e.g., no contextual support) may at the same time be simple in terms of another feature (i.e., have a clear structure). As Ellis (2003) noted, while principles gleaned from research can inform the design of task-based courses, the selection and grading of tasks will need to rely on the intuitions of experienced course designers.

Teaching Tip

- Sequence tasks in terms of their difficulty by drawing on the various factors that researchers have identified as influencing task complexity, but by also making use of your experience and intuition as to what will be the ‘right level’ of task for your students.

The research is perhaps of greater value when it comes to considering the organization of task-based lessons. There is general recognition that a task-based lesson can consist of three phases—a pretask phase, the main task phase, and the posttask phase (see, for example, Lee, 2000; Skehan, 1996; Willis, 1996). Of these three phases, only the main task phase is obligatory. In other words, a lesson can consist of all three phases, a pretask and main phase, a main-phase and follow-up phase, or just the main phase. To date the research has focused mainly on options relating to the pretask phase (e.g., pretask planning) and the main task phase (e.g., online planning and focus-on-form). Little attention has been paid to the posttask phase. This research suggests the value of manipulating both pretask and online planning conditions. It also points to the importance of incorporating a focus-on-form in the main task phase. In this respect the research contradicts the advice often given to teachers to avoid corrective feedback while performing a task. Ellis and Shintani (2014) found that popular teacher guides recommend that teachers should focus solely on ‘fluency’ when learners are performing tasks. Willis (1996) too advised teachers to “stand back and let the learners get on with the task on their own” (p. 54), a view that she continues to hold (see Willis & Willis, 2007). However, both theory and research point to the importance of attracting learners’ attention to form as they communicate as well as before and after performing a task.

Teaching Tips

- Don’t be frightened to focus learners’ attention on form while they are performing a task: quick ‘time-outs’ from communicating do not interfere with the communicative flow and they facilitate learning.
- Make use of input-based tasks with beginner learners. Also use an input-based task as a preparation for performing an output-based task.

Disappointingly, there are few studies of whole lessons involving a task. An interesting exception is Samuda’s (2001) study. Samuda was concerned with the role of the teacher in a

task-based lesson, arguing that teachers needed to find “ways of working *with* tasks to guide learners toward the types of language processing believed to support L2 development” (p. 120, emphasis in original). She documented a lesson based on a focused task designed to provide opportunities for learners to use epistemic modal verbs (e.g., *might* and *must*). She found that initially the learners avoided the use of these verbs opting instead to express degrees of probability by means of adverbs (e.g., ‘possibly’ and ‘probably’) and that when the attempt to scaffold their use interactionally failed, the teacher took time out from the performance of the task to provide a brief explicit explanation of the modal verbs and their meanings. This resulted in the learners then trying to use the verbs. This study is interesting because it shows—contrary to the claims of Swan and Sheen—that explicit grammar teaching can have a place in a task-based lesson.

Future Directions

It is likely that researchers will continue to investigate individual tasks, focusing on how their design features and method of implementation affect performance. We can expect that much of this research will be quantitative in nature with less reliance on traditional CAF measures and theory-driven improvements in how task performance is measured. We can also expect a more critical look at how task complexity affects the cognitive processing that occurs when a task is performed. Révész (2014), for example, has pointed to the importance of finding independent ways of investigating cognitive processing. We can also expect to see qualitative methods employed to examine both pretask options such as pretask planning and learners’ perceptions of tasks. There is also a clear need for more studies of complete task-based lessons (such as Samuda, 2001) and for longitudinal studies (such as Lambert & Robinson, 2014), which examine the implementation of a task-based materials in specific instructional contexts over time. To satisfy the critics of TBLT, more comparative studies (such as Shintani, 2015) are also needed.

To date there are no mainstream task-based course textbooks published by prominent publishers, although there are some such locally produced courses (e.g., Cutrone & Beh, 2015). The reluctance of major publishers to solicit and publish task-based courses is indicative of a resistance to task-based teaching in some teachers and teacher educators. In many instructional contexts TBLT constitutes an innovation and, as evaluations such as Carless (2004) have shown, teachers face problems in implementing TBLT. If TBLT is to become ‘mainstream,’ teachers will need training and support. There is a need to build up expertise about how this can be best provided, so studies reporting training programs such as Erlam (2016) are especially welcome.

References

- Anton, M., & DeCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review*, 54, 314–342.
- Baralt, M., Gilabert, R., & Robinson, P. (2014). *Task sequencing and instructed second language learning*. London: Bloomsbury.
- Breen, M. (1989). The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 187–206). Cambridge: Cambridge University Press.
- Carless, D. (2004). Issues in teachers’ reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly*, 38(4), 639–662.
- Cutrone, P., & Beh, S. (2015). *Welcome to Kyushu Japan: A task-based approach to EFL learning*. Tokyo: Shoakusha.

Rod Ellis

- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge: Cambridge University Press.
- Dewey, J. (1938). *Experience and education*. New York: Collier Books.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit knowledge. *Studies in Second Language Acquisition*, 27, 305–352.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? *Studies in Second Language Acquisition*, 24(2), 223–236.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). *Planning and task-performance in a second language*. Amsterdam: John Benjamins.
- Ellis, R. (2009a). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
- Ellis, R. (2009b). The differential effects of three types of planning on fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 30, 474–509.
- Ellis, R. (2010). Second language acquisition and language teaching materials. In N. Harwood (Ed.), *English language teaching materials: Theory and practice* (pp. 33–57). Cambridge: Cambridge University Press.
- Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London: Routledge.
- Erlam, R. (2016). I'm still not sure what a task is: Teachers designing language tasks. *Language Teaching Research*, 20, 279–299.
- Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18(3), 299–324.
- Gass, S., & Mackey, A. (2007). Input, interactions, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 180–206). Mahwah, NJ: Lawrence Erlbaum.
- Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.
- Howatt, A.P.R. (1984). *A history of English language teaching*. Oxford: Oxford University Press.
- Jackson, D., & Seuthanpronkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, 63, 330–367.
- Johnson, K. (1982). *Communicative syllabus design and methodology*. Oxford: Pergamon.
- Klippel, F. (1984). *Keep talking*. Cambridge: Cambridge University Press.
- Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics*, 35, 607–614.
- Lambert, C., & Robinson, P. (2014). Learning to perform narrative task: A semester long study of task sequencing effects. In M. Baralt, R. Gilabert, & P. Robinson (Eds.) *Task sequencing and instructed second language learning* (pp. 207–230). London: Bloomsbury.
- Lee, J. (2000). *Tasks and communicating in language classrooms*. Boston, MA: McGraw-Hill.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40, 243–249.
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, 361–386.
- Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–99). Clevedon: Multilingual Matters.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie & T.K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2: Second language acquisition* (pp. 413–468). New York: Academic.

- Long, M. (2015). *Second language acquisition and task-based language teaching*. New York: Wiley.
- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22(4), 471–497.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–453). Oxford: Oxford University Press.
- Pica, R., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language research and instruction. In S. Gass & G. Crookes (Eds.), *Task-based learning in a second language* (pp. 171–192). Clevedon: Multilingual Matters.
- Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 23–75). Clevedon: Multilingual Matters.
- Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
- Révész, A. (2014). Task-based learning: Research methods. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–6). Chichester: Wiley-Blackwell.
- Richards, J., & Rogers, T. (2014). *Approaches and methods in language teaching* (3th ed.). Cambridge: Cambridge University Press.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27–57.
- Robinson, P. (2011). Task-based language learning: A review of the issues. *Language Learning*, 61, 1–36.
- Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 379–400). Harlow: Longman.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Basingstoke: Palgrave Macmillan.
- Seedhouse, P. (2005). “Task” as research construct. *Language Learning*, 55(3), 533–570.
- Sheen, R. (2003). Focus on form—a myth in the making? *ELT Journal*, 57(3), 225–233.
- Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research*, 15, 137–158.
- Shintani, N. (2015). The incidental grammar acquisition in focus on form and focus on forms instruction for young, beginner learners. *TESOL Quarterly*, 49, 115–140.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 167–185). Harlow: Longman.
- Skehan, P. (Ed.). (2014). *Processing perspectives on task performance*. Amsterdam: John Benjamins.
- Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26(3), 376–401.
- Tong-Fredericks, C. (1984). Types of oral communication activities and the language they generate: A comparison. *System*, 12, 133–134.
- Van den Branden, K. (Ed.). (2006). *Task-based language education: From theory to practice*. Cambridge: Cambridge University Press.
- Widdowson, H. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: University Press.
- Willis, J. (1996). *A framework for task-based learning*. Harlow: Longman.

Cognitive-Interactionist Approaches to L2 Instruction

YouJin Kim

Background

Over the last few decades, the field of second language acquisition (SLA) has demonstrated extraordinary methodological, theoretical, and pedagogical development. As that research scope has expanded, a vital subfield called instructed second language acquisition (ISLA) has also emerged. According to Loewen (2015), ISLA is an academic discipline that studies “how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one’s first” (p. 2). Various approaches to SLA address different degrees of the relationship between L2 learning processes and instruction. For instance, Ortega (2015) states that while some theories address no relationship between L2 learning processes and instruction (e.g., Universal Grammar Theory), other theories make more specific proposals regarding how to promote L2 instruction by manipulating instructional conditions (e.g., Input Processing Theory). Among the various approaches to SLA, cognitive-interactionist approaches have made clear suggestions for optimal features of L2 instruction and contributed most notably to the development of ISLA.

The current chapter adopts Loewen’s definition of ISLA, discusses cognitive-interactionist approaches to L2 instruction in ISLA research, and highlights the roles of both native speaker–learner and learner–learner interaction. The roots of cognitive-interactionist approaches to SLA can be found in Long’s initial formulations of the Interaction Hypothesis (Long, 1981). Building on Hatch’s (1978) claims of the importance of carrying out conversations in language learning, the original version of Long’s Interaction Hypothesis posited that conversation features increase input comprehensibility for L2 learners. This view was highly influenced by various research trends such as Krashen’s Input Hypothesis (Krashen, 1980) and foreigner talk research (Ferguson, 1971). According to the Input Hypothesis, input that is slightly above a learner’s current interlanguage system is the main driving force for L2 learning (Krashen, 1985). This view of input motivates the belief that the input must be comprehended by the learner (i.e., $i + 1$) in order to assist the language acquisition process. In Long’s original version of the Interaction Hypothesis (Long, 1981), his belief of the role of input in L2 development concurred with Krashen’s hypothesis. However, Long expanded

this understanding by claiming that conversational modifications transform L2 input into comprehensible input and, as a result, facilitate L2 development.

In the 1980s, motivated by Krashen's Input Hypothesis and/or Long's Interaction Hypothesis, researchers examined how to make input more comprehensible to learners (e.g., Pica, Young, & Doughty, 1987). In particular, early interaction studies examined the occurrence of interactional modifications as outcome variables (for reviews see Mackey, 2012; Plonsky & Gass, 2011). As a result of this early work, two theoretical claims that arose during the 1980s and early 1990s pushed for a revision of Long's original Interaction Hypothesis: Schmidt's (1993) Noticing Hypothesis and Swain's (1985) Output Hypothesis. Schmidt's Noticing Hypothesis is subsumed in the revised version of the Interaction Hypothesis. According to Schmidt, noticing of forms is necessary and sufficient for L2 learning. Swain (1985), based on her research in Canadian immersion contexts, claimed that while comprehensible input is necessary for L2 development, on its own, it is insufficient. She argued that language production is necessary because it forces learners to move beyond semantic processing of language during comprehension and focuses their attention toward syntactic use of language. According to Swain's Output Hypothesis, producing output thus plays a critical role in L2 acquisition because it encourages learners to notice gaps in their interlanguage system (i.e., noticing), gives learners a chance to test their linguistic hypotheses (i.e., hypothesis testing), and fosters the co-construction of knowledge when learners use language to reflect on language use (i.e., metalinguistic awareness) (Swain, 1995). Drawing upon these theories, Long's updated Interaction Hypothesis (1996) states: "Negotiation for meaning, and especially negotiation that triggers interactional adjustments by the native speaker or more competence interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways" (p. 451).

Since the proposal of the updated Interaction Hypothesis, the field of SLA has witnessed an explosion of empirical studies with the overarching goal of investigating how interaction facilitates L2 learning. Moving beyond the studies that focus on the occurrence of interactional features in various conditions from the early 1990s, researchers began providing evidence of the benefits of interaction on L2 development (e.g., Ellis, Tanaka, & Yamazaki, 1994; Gass & Varonis, 1994). A seminal study by Gass and Varonis (1994) showed a direct relationship between interaction and learners' subsequent linguistic production. Drawing on interaction data produced by 16 native speaker–nonnative speaker pairs, Gass and Varonis found a positive role for negotiation of meaning on learners' delayed L2 production and comprehension. Since then, using more rigorous research designs, studies have advanced theoretical claims related to interactional features, such as corrective feedback (i.e., response to learner errors) and modified output (i.e., modification of the original erroneous utterance) (McDonough, 2005). For instance, Mackey (1999) used a pretest–posttest research design to examine whether conversational interaction facilitated the acquisition of English question formation. Her findings indicated that learners who actively participated in interaction led to a greater language development (i.e., production of syntactically advanced question forms) compared to those who simply observed interaction without participating in it or who carried out tasks using scripted interaction.

Today, many studies later, we have ample evidence that interaction does foster L2 development, as summarized in several meta-analyses (e.g., Lyster & Saito, 2010; Plonsky & Gass, 2011). For instance, Keck, Iberri-Shea, Tracy-Ventura, and Wa-Mbaleka (2006) and Mackey and Goo (2007) showed medium-to-large effects for interaction. Based on the clear empirical support for the benefits of interaction, Gass and Mackey (2007) claimed that "it is now commonly accepted within the SLA literature that there is a robust connection between

interaction and learning” (p. 176). Therefore, as a field, there is agreement that what was once viewed as a hypothesis (i.e., Interaction Hypothesis) can now be viewed as an approach to SLA (Mackey, Abbuhl, & Gass, 2013).

The major constructs of cognitive-interactionist approaches to L2 learning are reviewed under the following key concepts. In sum, cognitive-interactionist research to SLA has strived to account for acquisition by examining the input that learners receive, the interaction that they engage in, and the output that they produce (Gass & Mackey, 2007). These also include the most fine-grained investigated constructs, such as noticing, focus on form, and corrective feedback.

Key Concepts

- **Input:** The language that is available to a learner through any medium (e.g., listening, reading). It provides positive evidence.
- **Interaction:** Conversation in which the learners engage. It can be carried out in person, online, or through other mobile-mediated communication settings.
- **Output:** The oral or written language that is produced by learners.
- **Noticing:** Paying attention to linguistic input with some level of awareness.
- **Corrective feedback:** A written or oral response to learners’ errors. It provides negative evidence (i.e., information about what is not possible in a target language).
- **Focus on form:** Spontaneous attention to linguistic forms during meaning-oriented activities.
- **Modified output:** Learners’ response to feedback that is more target-like than the original utterance.

Current Issues

In the present chapter, interaction is discussed in terms of native speaker–learner and learner–learner interaction. One of the most widely examined topics in both native speaker–learner and learner–learner interaction research domains is focus on form. In the current article, it is defined as—spontaneous attention to linguistic forms during meaning-oriented activities. Among different types of focus on form techniques (e.g., input enhancement, corrective feedback), a large amount of research focuses on corrective feedback. This emphasis is due to both theoretical and pedagogical motivations (see Loewen, 2012, for review). From a theoretical perspective, an interactionist approach to SLA values the role of corrective feedback in language development because it promotes noticing of linguistic forms in unobtrusive ways. Pedagogically speaking, it addresses long-standing concerns related to a lack of accuracy development in communicative language teaching. Recent development in interaction research involves exploring intervening variables that affect the role of corrective feedback such as feedback types (e.g., recasts vs. elicitation; Nassaji, 2009), individual differences (e.g., working memory, aptitude, anxiety, proficiency; Li, 2013), and target linguistic features (e.g., questions, passives; Mackey, 2006). For instance, working memory has been increasingly examined, and research has shown that individual differences in working memory capacity is associated with the role of corrective feedback in language learning (e.g., Goo, 2012; Li, 2013; Mackey, Philp, Fujii, Egi, & Tatsumi, 2002; Mackey & Sachs, 2012).

In a similar vein, learner–learner interaction studies have also focused on examining how to facilitate negotiation of meaning and focus on form opportunities during meaning-oriented

discourse. A major development in this research domain comes from the use of interactive tasks (Ellis, 2003; Plonsky & Kim, 2016). For learner–learner interaction, interactive tasks have widely been used to either elicit interactional features or explore task effects on interaction-driven learning. As tasks have received increased attention among both task-based language teaching (TBLT) and ISLA researchers, there has been an accompanying surge of research examining learners’ task performance in terms of interactional features during collaborative tasks (see Kim, 2015 for a review). Recent studies have examined task design and implementation variables such as task complexity, task repetition, and task planning time.

Finally, due to the development of instructional technology, computer-mediated communication (CMC) has been increasingly explored (see Sauro, 2011 and Ziegler, 2016). Researchers have shown the helpful features of face-to-face interaction also taking place in synchronous computer-mediated communication (SCMC) (see Sauro, 2011 for a review). Specifically, learners are provided with opportunities to interact, produce language, and modify their output in response to any communication difficulties, as well as respond to feedback from an interlocutor in an authentic communicative setting. Moreover, some studies have indicated that SCMC can provide learners with advantages over face-to-face interactions including increased opportunities for learners’ attention to be drawn to the form of the language, and more time for them to understand and process what they hear and see (Ziegler, 2016).

Empirical Evidence

Native Speaker-Learner Interaction: Features of Corrective Feedback

Within the field of ISLA, teachers’ use of focus on form techniques, particularly corrective feedback practices, has been the focus of numerous empirical studies (see Brown, 2016; Li, 2010; Lyster & Saito, 2010; Lyster, Saito, & Sato, 2013; Mackey & Goo, 2007; Nassaji, 2016; Russell & Spada, 2006 for meta-analysis or synthesis reports on corrective feedback). With the growing preference for communicative language teaching approaches in L2 pedagogical contexts, several questions persist: how can students’ attention be drawn to linguistic forms, and how can accuracy be developed while working on fluency? Loewen (2012, p. 26) concisely presents five core issues that were examined in previous corrective feedback studies: (1) Does feedback occur naturally in the L2 classroom? (2) What are the characteristics of naturally occurring feedback? (3) Is feedback effective for L2 learning? (4) What characteristics of feedback influence its effectiveness? (5) What contextual characteristics of feedback influence its effectiveness?

First, early classroom-based corrective feedback studies were interested in examining whether feedback occurs naturally in diverse L2 classroom contexts. Important observations derived from earlier descriptive studies revealed that corrective feedback naturally occurs in language classrooms. However, such feedback does not occur with similar frequency across instructional contexts (e.g., Lyster & Mori, 2006). Furthermore, as Loewen’s second core point suggests, these descriptive studies also identified various types of corrective feedback, including recasts, clarification requests, confirmation checks, elicitation, and metalinguistic feedback, which all occurred at various frequencies. These various feedback types have also been described along a continuum that ranges from implicit to explicit (e.g., Lyster & Saito, 2010; Lyster et al., 2013). They can also be classified as input-providing feedback (e.g., recasts) or output promoting feedback (e.g., elicitation) (Ellis, 2008). In a recent meta-analysis, Brown (2016) examined the proportion of corrective feedback types teachers provide in L2 classrooms based on 28 classroom-based corrective feedback studies. Brown included

85 teachers across 11 countries, as well as seven target languages. A total of 7,188 corrective moves were tallied, and the findings showed that reformulations (recast and explicit correction) outweighed prompts (66% vs. 30%).

Loewen's third core issue addresses the relationship between corrective feedback and L2 learning. Learners' responses to different types of feedback (i.e., uptake, modified output) and/or scores on posttests were used as a way to investigate the effects of corrective feedback (Loewen & Nabei, 2007; Nassaji, 2009). To date, various types of measures have been used, including production tests for target structures (McDonough, 2005) and grammaticality judgment tests that examine learners' ability to identify grammatical and ungrammatical sentences (e.g., Loewen & Nabei, 2007). Finally, in order to address a direct relationship between corrective feedback and language learning, tailor-made posttests (i.e., test items that are designed based on the focus of feedback) have also been implemented (Loewen & Philp, 2006). Brown (2016) reported that 42.7% of feedback episodes identified in classroom-based corrective feedback studies targeted grammar, whereas 27.6% and 22.4% of the feedback episodes targeted lexis and phonology, respectively.

Building on the positive findings of corrective feedback, researchers have examined how different feedback characteristics influence the effectiveness of feedback in relation to noticing and/or L2 development. In a number of quasi-experimental studies, researchers have compared different types of feedback in terms of the degree of L2 learning, and recasts have often been compared to other types of feedback. For instance, recasts were as effective as prompts for young ESL learners with high pretest scores but less effective than prompts for learners with low pretest scores (Ammar & Spada, 2006). Additionally, Lyster and Saito (2010) noted that classroom learners benefit from the positive evidence in recasts as well as negative evidence, but may benefit even more from the negative evidence in prompts that create greater demand for producing modified output.

Overall, Lyster et al. (2013) claim that oral corrective feedback is significantly more effective than no corrective feedback, and also, that prompts or explicit correction tends to show more learning gains than recasts. However, because even one type of corrective feedback might not be consistent between studies and contexts (e.g., various degrees of explicitness), it is difficult to confirm the degree of effectiveness among various types of corrective feedback. What seems to be necessary is to understand what features of corrective feedback contribute to the effectiveness of corrective feedback. For instance, Loewen and Philp (2006) examined the provision and effectiveness of recasts in 12 adult ESL classrooms during 17 hours of meaning-based interaction. They analyzed the linguistic focus, length, prosodic emphasis, segmentation, number of changes, and intonation of the recasts. Based on a tailor-made posttest, the findings indicated that recasts were beneficial at least 50% of the time. Stress, declarative intonation, one focal linguistic change within a recast, and multiple feedback moves were predictive of successful uptake. Furthermore, interrogative intonation, shortened length, and one focal linguistic change were predictive of the development of accuracy.

Other studies have shown that the effectiveness of feedback may vary depending on the type of linguistic target (Jeon, 2007; Mackey, 2006). For example, Mackey (2006) showed that feedback targeting question forms were noticed more than those targeting past tense morphemes in ESL classroom. Jeon (2007) examined Korean language learners' interaction-driven language learning, and found that corrective feedback promotes L2 learning of nouns, verbs, and object relative clauses more effectively than Korean honorific agreement morphology.

For his last core point, Loewen (2012) highlights that different contextual features also affect the effectiveness of feedback (e.g., instructional variables, interlocutor variables). In

terms of the role of contextual characteristics surrounding the provision of feedback, instructional variables such as instructional contexts and teaching experiences are also relevant. For instance, immersion contexts have been shown to elicit a smaller amount of feedback than language classes (e.g., Ellis, Basturkmen, & Loewen, 2001; Lyster & Ranta, 1997; Sheen, 2004). Zyzik and Polio (2008) also reported a near absence of corrective feedback in content-based university Spanish classes. Learners' education level as an instructional contextual variable was found to be an important factor as well. In his meta-analysis, Brown (2016) reported that adults received a significantly greater proportion of recasts than high school students, and elementary-level students received a similar rate of recasts/prompts as adult. Furthermore, younger learners received a significantly greater proportion of corrective feedback focusing on lexis compared to adults, while adults received a greater proportion of corrective feedback on pronunciation compared to elementary students. Analysis of second language and foreign language teaching contexts suggested that second language teachers targeted phonological errors significantly more than foreign language teachers. Lexical errors were targeted more consistently between contexts, while grammar was addressed more often in foreign language contexts compared to second language contexts. Brown (2016) also found that among various teacher variables (e.g., native vs. nonnative), teaching experience and education/training was found to moderate corrective feedback choices. For instance, Junqueira and Kim (2013) compared the corrective feedback practices of a novice and a more experienced teacher in oral communication classes. In that study, the more experienced teacher provided more corrective feedback, and targeted more types of linguistic features.

When discussing corrective feedback, learners' response to feedback in a form of uptake or modified output has been widely examined. For instance, McDonough (2005) compared the effects of four conditions on the development of English questions: enhanced opportunities to modify, opportunities to modify, feedback without opportunity to modify, and no feedback. McDonough found that students who had modified output opportunities showed a greater degree of learning compared to those without modified output opportunities. Her findings suggest that receiving solely corrective feedback might not be sufficient. Rather, language development through interaction is contingent upon learners' ability to notice the gap between their interlanguage and the corrective feedback, as well as the production of modified output. Recently, Gurzynski-Weiss and Baralt (2015) have expanded our understanding of the role of modified output in interaction-driven language learning. Their findings suggest that after feedback, partial modified output (i.e., learners isolated and repeated only the element that had been corrected in feedback) was the greatest predictor of accurate noticing of feedback in both face-to-face and text chat interaction settings.

Additionally, individual difference variables such as working memory (Mackey et al., 2002), language aptitude (Li, 2013), and language anxiety (Sheen, 2008) impact the effectiveness of corrective feedback. Among various factors, working memory has been increasingly investigated in corrective feedback studies, and has addressed a complex picture of the role of corrective feedback in L2 learning. For instance, Mackey et al. (2002) suggest that working memory was positively associated with the noticing of recasts, and Mackey, Adams, Stafford, and Winke (2010) also showed that working memory is positively correlated with the amount of modified output during collaborative tasks. Mackey and Sachs (2012) noted that older learners with higher working memory demonstrate question development through interactive tasks. Additionally, Goo (2012) revealed that while recasts and metalinguistic explanations were equally effective on learners' acquisition of that-trace filter, working memory significantly mediated the effectiveness of recasts, suggesting that

YouJin Kim

executive attention is involved in the noticing of recasts. However, Li (2013) found that working memory mediated the effects of explicit feedback rather than recasts. These differences might be due to the nature of target structure, however. For example, while Goo targeted English language acquisition, Li focused on Chinese language acquisition. In sum, over the last three decades, research on corrective feedback as a way to provide focus on form opportunities has contributed to our understanding of the role of interaction in language learning from both theoretical and pedagogical perspectives.

Learner–Learner Interaction in Classroom Contexts: Task Design and Implementation Variables

Within interaction-based research, the domain of learner–learner interaction research has been developed noticeably over the last two decades (e.g., Philp, Adams, & Iwashita, 2014; Sato & Ballinger, 2016). Building on the notion that input, interaction and output opportunities are necessary conditions for L2 development, pair and group work is widely implemented in foreign and second language classrooms. Researchers and language practitioners are highly interested in identifying what may facilitate language learning through learner–learner interaction. The development of learner–learner interaction studies has been accompanied by task-based research. Accordingly, there has been significant attention to task-based interaction in ISLA literature. Plonsky and Kim (2016) analyzed 85 task-based learner performance studies, and found that interaction features were the second most frequently analyzed feature. Among interactional features, language-related episodes (LREs; instances in the interactions where learners talk about, question, and/or self-or-other correct language use: Swain & Lapkin, 1998) were the most widely analyzed interactional features in task-based studies. Such findings show the close interface between task-based research and interaction studies in the field of ISLA.

From a cognitive-interactionist perspective, the relationship between task-related variables (e.g., task design and implementation) and the occurrence of interactional features, as well as the subsequent learning outcomes, is of particular interest to researchers. Some common concerns regarding learner–learner interaction are related to learners' adoption of other learners' errors during learner–learner interaction. Furthermore, within the domain of ISLA, how to promote beneficial interactional features during learner–learner interaction is of particular interest. In order to address these concerns, researchers have analyzed interactional features, and addressed the resolution outcomes of language-related discussion.

Targeted interactional features during learner–learner interaction include LREs, form-focused episodes, the provision of, noticing of, and use of corrective feedback, and negotiation of meaning (e.g., McDonough, 2004; Lyster et al., 2013; Philp, Oliver, & Mackey, 2006; Révész, 2011). Some studies have established a positive relationship between interactional features and the subsequent language learning as a result of task-based interaction in the classroom (e.g., Adams, 2007; Newton, 2013). These studies focused on lexical items as well as grammatical forms such as English question formation, past tense, and prepositions (e.g., Kim, 2012; Nuevo, 2006; Patanasorn, 2010).

The underlying assumptions of cognitive-interactionists is that task type, task design, and implementation features might manipulate learner cognitive processes and in turn affect learners' task performance and subsequent language development. One task design variable that has received a great deal of recent attention is task complexity. Robinson's (2001) Cognition

Hypothesis predicts that increasing complexity along with resource directing variables (e.g., the presence of reasoning demands) will promote more interactional features such as negotiation for meaning and corrective feedback, which in turn facilitate interaction-driven language development. Studies that examine task complexity have operationalized interactional features or learning opportunities as LREs and various feedback types. Previous research that was conducted during learner–learner interaction in classroom contexts, has partially supported the benefits of carrying out more cognitively demanding tasks on interaction-driven learning opportunities, especially when learning opportunities were operationalized as LREs. However, outcomes may depend on the mediating role of task design and learner factors (e.g., proficiency). For instance, Kim (2009) showed that learner proficiency and task types mediate task complexity effects. The findings suggest that with more proficient learners, the complex version of the picture narration task elicited more LREs than the simple version. However, during the picture difference task, such pattern was found with only lower level learners. Although no significant differences were found in resolution of LREs, more complex tasks tended to draw a slightly higher number of correctly resolved LREs. Additionally, Révész (2011) examined the role of task complexity in interaction-driven learning opportunities during decision-making tasks with ESL learners. She found that the more complex task caused greater amounts of LREs during learner–learner interaction, but the amount of different types of corrective feedback did not support the Cognition Hypothesis as there was no significant increase in the amount of corrective feedback with complex tasks compared to simple tasks.

In terms of the relationship among task complexity, interactional features, and L2 development, there has been mixed findings. While Nuevo (2006) indicated no significant association between task complexity and L2 learning in adult ESL classrooms, Kim (2012) found a greater number of LREs during complex tasks, which in turn facilitated Korean EFL learners' question development. More recently, Kim and Taguchi (2015) provided further evidence for the long-term benefits of carrying out more cognitively demanding tasks in terms of learning request-making expressions among Korean EFL adolescents. Kim and Taguchi expanded the scope of task complexity studies by focusing on pragmatics as a target linguistic area, and discussing long-term benefits of learner–learner interaction for learning request-making expressions.

Researchers have also been increasingly interested in how task implementation factors in the classroom differentially affect interaction-driven language learning. The two most widely addressed factors in previous classroom-based studies are task planning and task repetition. It was claimed that due to L2 learners' limited attentional resources (Skehan & Foster, 2005), pretask planning and task repetition would enable learners to produce higher quality language samples (i.e., fluency, accuracy, complexity) as well as to pay more attention to linguistic forms during task-based interaction (e.g., Foster & Skehan, 1999; Mehnert, 1998; Ortega, 1999; Sangarun, 2005; Skehan & Foster, 2005; Truong & Storch, 2007). Many task planning and task repetition studies have implemented monologic tasks, and only a few studies were conducted using collaborative tasks from an interactionist perspective. For instance, Philp et al. (2006) investigated how the amount of planning time impacts the amount of interaction. They found that little or no planning time led to more talk and the increasing amount of feedback provision between ESL learners. Truong and Storch (2007) examined group planning prior to oral presentation tasks, and found that groups composed of mixed proficiency learners were more interactive and focused on both content and language issues related to the upcoming presentation tasks during task planning.

Previous planning studies have also addressed the role of guided planning in task performance. Park (2010) compared the effects of task planning with and without specific instructions regarding linguistic aspects on task performance. Korean EFL learners were asked to pay attention to content, organization, vocabulary, and to grammar during collaborative narrative tasks. The findings indicated that, regardless of pretask instructions and planning opportunity, the learners' attention was oriented toward vocabulary. Additionally, Kim (2013a) investigated the role of showing task modeling videos during planning time, and the findings suggested that the use of task modeling videos during planning time might facilitate interaction and collaborative task performance. Despite a growing amount of research, the role of guided planning has not been conclusive, and this might be due to different techniques that the instructors use during planning time (e.g., using grammar review handouts, showing task modeling videos).

A second task implementation factor that has received a growing attention is task repetition. Bygate (2001) claimed that task repetition would allow learners to allocate more cognitive resources to language rather than task content. Such process is believed to promote interlanguage development. Similar to other task design and implementation variables, task repetition has been mostly investigated using monologic oral tasks; and only recently have task repetition studies investigated the effect of task repetition during task-based learner–learner interaction with intact L2 classrooms (Azkarai & García Mayo, in press; Kim, 2013b; Patanasorn, 2010). These studies also suggest different ways in which task repetition can be operationalized: repeating exactly the same tasks, repeating procedure only, and repeating content only. Patanasorn (2010) investigated the effects of different characteristics of task repetition in the acquisition of past tense morphology during learner–learner interaction in Thai EFL university contexts. Results showed that repeating the same task procedure was more beneficial for promoting the development of past tense accuracy, and that content repetition was more beneficial for global fluency at the expense of past simple accuracy. Kim (2013b) compared the impact of task repetition (i.e., exact repetition) and procedural repetition on Korean EFL learners' production of LREs (i.e., attention to linguistic form) during collaborative tasks. The results showed the benefits of repeating the same task procedure with different contents (i.e., procedural repetition condition) on the production of LREs.

Another recently developed topic within the research domain of interactive task design is interactive alignment (or priming), the phenomenon of speakers' tendency to use linguistic structures that they have recently heard. During interaction, there is ample evidence that speakers are likely to converge toward similar linguistic patterns and constructions such that individuals reuse similar expressions, grammatical structures, and patterns of pronunciation previously employed by their interlocutor (McDonough & Trofimovich, 2009; Trofimovich, 2016).

Interaction researchers have adopted priming mechanisms to task design as priming may prompt linguistic convergence during interaction (see McDonough & Trofimovich, 2009 for a review). Three types of priming have been introduced in the field of SLA: structural (also called syntactic), auditory, and semantic priming (McDonough & Trofimovich, 2009). The premise of structural priming is that although alternate structures that express similar meaning are available to speakers, they will most likely produce a syntactic construction that they have just been exposed to in the preceding discourse. For example, during a conversation, one speaker might produce a passive construction (*The letter was delivered by my uncle*). The same speaker or his/her interlocutor is likely to produce another passive construction later in the same conversation (*John's computer was fixed by Tom yesterday*) rather than an active construction (*Tom fixed John's computer yesterday*) which is an alternative form.

In order to facilitate interaction-driven language learning between learners, researchers have built structural priming techniques into task design. For instance, McDonough and her colleagues have conducted a series of classroom-based studies that explored the role of priming tasks on the elicitation of target structures and learning outcomes. During collaborative structural priming tasks, one interlocutor has a list of sentence models (i.e., primes) which include target structure (e.g., passives, phrasal verbs). The other student is asked to produce language output using a prompt, which is often a single word or a phrase (e.g., McDonough & Chaikitmongkol, 2010). The occurrence of structural priming is demonstrated when the speaker with a prompt (e.g., given verbs and nouns) produces an utterance that has the target structure that was the structure of the preceding prime. McDonough, Trofimovich, and Neumann (2015) implemented collaborative structural priming tasks in an English-for-academic-purposes class over a 13-week semester targeting three English structures: passives, relative clauses, and adverbial clauses. The findings suggested that priming tasks facilitated the production of relative target structures and adverbial clauses. However, no difference was found between priming and no priming conditions in the production of passives.

Recently, auditory priming has received some attention in ISLA research. Auditory priming refers to speakers' tendency to process a spoken word more quickly and to produce a word more accurately when they have previously heard that word compared to a novel word (McDonough & Trofimovich, 2012; Trofimovich, 2016). Trofimovich, McDonough, and Foote (2014) examined the effects of auditory priming during collaborative learner–learner interaction tasks in ESL classroom contexts. The target words with target stress patterns (e.g., intelligent 4–2 stress pattern; four-syllable word with the stress on the second syllable) were embedded into four collaborative, information-exchange tasks that were created as a part of the regular course materials. For each of these tasks, each learner was given sentences that included words with target stress patterns. Based on the accuracy rates of target stress patterns, Trofimovich et al. found that auditory priming occurred during collaborative tasks. In terms of the occurrence of priming, research has suggested that multiple language features can be primed simultaneously during collaborative tasks. For instance, Trofimovich, McDonough, and Neumann (2013) revealed that integrated auditory (i.e., stress) and structural primes were more successful at eliciting the target forms, compared to auditory-only and structure-only primes. In sum, priming mechanism could be used in designing collaborative tasks to facilitate the production of target linguistic features during interaction.

Technology-Mediated Interaction: Synchronous computer-mediated communication (SCMC)

Another notable development in cognitive-interactionist research is the expansion of interactional contexts. The use of diverse instructional technology in educational contexts and the creation of online courses have contributed to the expansion of interaction research (see McDonough & Mackey, 2013). Besides face-to-face interaction in classroom contexts, computer-mediated communication (CMC) interaction settings have received increasing attention in the interactionist research domain (Ziegler, 2016). In particular, SCMC has been attracting much attention in the SLA literature for its purported benefits for L2 development (Smith, 2005). SCMC refers to real-time interaction between people over a computer network (Stockwell, 2010), and it can be oral (voice chat) or written (text chat) modes. Researchers have explored the extent to which interactional features that occur during face-to-face

interaction take place in SCMC (see Sauro, 2011 for a review). The current review focuses on written SCMC (i.e., text chat).

In SCMC, unlike face-to-face interaction, learners have more time to understand and process what they see or hear. Some researchers have therefore suggested that SCMC may provide learners with advantages over face-to-face interactions, including increased opportunities for learners' attention to the linguistic forms (see Ziegler, 2016 for review). Previous studies have demonstrated that unique features of SCMC (e.g., text-based and computer-mediated interaction, slower rate of interaction, little competition over turn-taking) may allow for advantages within SCMC contexts over face-to-face interaction contexts. In her recent meta-analysis, Ziegler (2016) provide detailed synthesis of SCMC interaction studies, and highlight the benefits of interaction in both FTF and SCMC. Earlier SCMC research explored whether interactional patterns that are often observed in FTF also happen in SCMC contexts, and the findings suggest that the SCMC mode provides the opportunities of interaction, negotiation for meaning, and feedback as in the FTF mode between learners (e.g., Pellettieri, 2000; Smith, 2003) and between learners and native speakers (e.g., Iwasaki & Oliver, 2003; O'Rourke, 2005). With regard to comparing SCMC and FTF in terms of the benefits of L2 learning, Ziegler found that there was a small effect size in favor of SCMC for overall L2 learning outcomes. Her findings also suggest that SCMC tended to favor the development of written skills, and the FTF condition slightly benefit more for the oral skills, which can be accounted for by DeKeyser's claim (2015) that practice and learning might be skill specific.

In terms of the role of interaction in SCMC contexts, researchers have investigated whether SCMC may promote noticing of target features compared to FTF interaction (see Sauro, 2011 for a review). A growing body of research has claimed that SCMC may have the great potential for promoting noticing of linguistic forms as compared to FTF interaction (e.g., Lai & Zhao, 2006). First, compared to FTF mode, SCMC allows for a slower pace of conversation and turn-taking during which learners may take longer processing time in comprehending and producing the target language. Additionally, learners can easily access previous chat messages throughout the conversational exchanges (Yuksel & Inan, 2014).

Several studies have compared the amount of negotiation of meaning between face-to-face interactions and written SCMC, and found that FTF interactions elicit more collaborative interaction than written SCMC mode (e.g., Fernández-García & Martínez Arbelaiz, 2003; Loewen & Wolff, 2016). Loewen and Wolff suggest that the slower pace of interaction in written SCMC mode might allow learners to take more time to monitor their language production, thus they may not need to be engaged in negotiation of meaning as much as in oral mode. Another interesting finding was that the frequency of different types of interactional features between the two modes was varied. For instance, the most common interaction features in the oral mode were confirmation checks and LREs, yet they were hardly occurred in the written SCMC mode.

Previous research has examined learners' noticing of linguistic forms in the form of LREs and form-focused episodes (FFE). So far the findings related to the occurrence of LREs and FFEs between the two modes have been mixed. For instance, Loewen and Reissner (2009) compared incidental focus on form in L2 classroom and chatroom by analyzing the occurrence of FFEs, in which learners pause their interaction and to focus on language features. Their findings showed that although both FTF and SCMC modes elicited FFEs, there were notably more FFEs in FTF conditions than SCMC conditions. In an attempt to examine a direct link between noticing and L2 acquisition, Shekary and Tahririan (2006) analyzed

text-based interactions between pairs of Persian learners of English using LREs as a unit of analysis to investigate whether linguistic forms can be acquired when noticed in LREs. Although they did not compare FTF and SCMC modes in terms of the amount of LREs, they compared the ratio of LREs to previous FTF studies, and suggested that the ratio of LREs to amount of talk in their study exceeded those in Williams's study (1999). They also provided further evidence of language learning after SCMC interaction by carrying out tailor-made posttests. The findings suggest that learners were able to remember the targeted linguistic items almost 70.3% of the time on the immediate posttest and 56.7% of the time on the delayed posttests. Based on these findings, Shekary and Tahririan suggested that noticing in text-based SCMC leads to acquisition.

Previous studies also examined the occurrence of modified output in different modalities. Recently, Gurzynski-Weiss and Baralt (2014) examined whether learners accurately noticed different types of feedback (e.g., recast, negotiation of meaning, clarification, repetition) in relation to error types (morphosyntactic, lexical, phonological, semantic, and spelling). They also compared learners' production of modified output in SCMC and FTF modes. The findings indicated that the learners perceived feedback the majority of the time in both modes and their perception was the most accurate for the lexical and semantic target, followed by morphosyntax in both modes. However, the findings indicated that modality did not differentially mediate learners' noticing of feedback, even though there was a significant effect of modality on the number of opportunities provided to learners for modifying output. Specifically, learners had significantly more opportunities to modify output in the FTF interaction, particularly after receiving feedback addressing lexical and morphosyntactic errors.

Similar to FTF interaction research, SCMC research has examined a variety of task factors and individual differences. For instance, Baralt, Gurzynski-Weiss and Kim (2016) compared the amount of different aspects of engagement (affective engagement, social engagement, and cognitive engagement) between the simple and complex groups in both FTF and written SCMC modes. They found that although carrying out more complex tasks in FTF mode facilitated learners' cognitive engagement, a low level of engagement was found in SCMC mode regardless of task complexity. Among different learner variables, anxiety and working memory have been widely examined. Because SCMC interaction allows for a self-correction before sending messages, researchers have claimed that learners might be less intimidated. For instance, Redmon and Burger (2004) demonstrated that students found engaging with classmates in online interaction less intimidating and less likely to be dominated by a single student than in a regular FTF classroom. This may indicate that learners' anxiety was greatly diminished during SCMC. On the contrary, Baralt and Gurzynski-Weiss (2011) found that learner state anxiety was not differentially mediated by two different modalities, SCMC and FTF interaction. Unlike common assumption in the literature, anxiety did not turn out significantly lower in the SCMC modality compared to FTF mode, even though learners spent an average of double the amount of time interacting via SCMC. Interestingly, this comparable level of anxiety between the two modes was not found to correlate with learners' responses about the task preference and in the background questionnaire.

Working memory as a cognitive learner variable has also been a focus of SCMC studies. With the increased use of technology for L2 learning, researchers have suggested that features and affordances of communication technologies may either reduce the burden on or induce a more effective use of working memory. Based on Levelt's speech production model and working memory theory, Payne and Whitney (2002) compared FTF interaction and SCMC interaction classes (those who conducted half of their class time in the

chatroom) in terms of their oral proficiency development whether working memory capacity predicts the rate of L2 oral proficiency development. The findings show that SCMC group outperformed the FTF group in the development of oral proficiency, which suggest that learners were able to develop their conversation skills through text chat. The results also show that written SCMC mode benefitted learners with lower working memory. In a follow-up study, Payne and Ross (2005) expanded their previous study by including discourse and corpus analytic techniques to explore how working memory capacity may affect the occurrence of repetition and other patterns of interactive features during chat sessions and oral proficiency development. The findings show that learners with low phonological working memory capacity produced a greater number of words per utterance than those with high phonological working memory capacity. Furthermore, students with high phonological working memory showed significantly more gains in their oral proficiency than the low-span students.

Overall, many researchers have claimed that it is not ideal to compare the characteristics and learning outcomes of FTF and SCMC interaction, nor to claim that one mode promotes X better than the other mode. Because these two modes originally involve different conditions from cognitive, affective, and social perspectives, it is not surprising that the research findings demonstrate differences between the two modes. As shown in the review, two modes are distinct from each other, and future research needs to treat both modes as different instructional contexts without making any generalizations.

Pedagogical Implications

Empirical research reviewed in the current chapter provides a variety of pedagogical implications. First, with regards to teachers' use of focus on form techniques, particularly corrective feedback, teachers should ideally include a variety of feedback, because no one type has been identified as the most effective type. Furthermore, as shown in Loewen and Philp (2006) with recasts, teachers might want to consider how to make such feedback more efficient and more salient to the learners. For instance, it is not wise to target multiple linguistic features in a single recast. Teachers might also want to manipulate intonation to make the recast more salient. Teacher trainers could also discuss explicitly the different focus on form techniques with student teachers, so that they recognize the benefits of these instructional practices and increase their ability to notice such practices when they arise during classroom interaction.

In terms of learner–learner interaction, we have strong evidence that it is beneficial in classrooms. What seems to be the most important at the present stage is to train learners to become autonomous interaction participants, given that recent training studies demonstrate the benefits of training (Fujii, Ziegler, & Mackey, 2016; Kim, 2013a; Sato & Lyster, 2012). Based on the review of the studies from the cognitive-interactionist perspectives, it is clear that task design features can also impact the degree of interaction-driven language learning. Tasks need to be designed to encourage learners to use language in meaningful contexts, to engage with target language, and to facilitate interaction. For instance, when task complexity is increased in a way that it encourages learners to use target features, learners might pay more attention to target linguistic features while carrying out collaborative tasks (Kim, 2012). Also, the use of specific techniques as a part of guided planning seems to encourage learners to pay attention to target forms during interaction. Finally, encouraging learners to work collaboratively to identify gaps and actively work to identify solutions for their LREs or FFEs may be helpful for L2 development.

Teaching Tips

- Teachers need to make corrective feedback salient as long as it does not interrupt natural conversation flows in class.
- Collaborative tasks need to be designed in a way that they elicit beneficial interactional features. Some factors that teachers should focus on are task complexity, task types, and task implementation (e.g., task repetition).
- Teachers want to recycle the procedure of entire tasks and/or parts of tasks with different content to increase learners' autonomy as task participants.
- Teachers need to provide training that focuses on how to participate in interactive tasks effectively for their learning. For instance, teachers can provide appropriate task modeling that demonstrates how to perform collaborative tasks and how to actively remain involved in learner–learner interaction attending to both form and meaning.
- Teachers might want to incorporate priming mechanisms in their task design so that some linguistic models are provided naturally.

Future Directions

Since the 1980s, interaction research in SLA has maintained its dynamic research agenda in terms of developing theory and research methodology (Mackey et al., 2013; Plonsky & Gass, 2011). Furthermore, it has offered a number of pedagogical implications in diverse instructional contexts. As Gass and Mackey (2015) claim, based on a significant amount of empirical work supporting the benefits of interaction in L2 learning, it is now referred to as the interaction approach to language teaching. Although considerable research has provided convincing evidence for the benefits of interaction in the L2 classroom, much more future research has yet to be done in the field.

For example, as one of the goals of ISLA research is to inform classroom instruction, it is necessary to connect interaction research and teacher training. For instance, Vásquez and Harvey (2010) had their SLA student participants conduct small-scale research projects in which they partially replicated Lyster and Ranta's (1997) study in their ESL classes. The findings indicated an important shift in the graduate students' beliefs about corrective feedback after the research projects, from primary concern with the affective dimension and face-threatening nature of corrective feedback to a greater understanding and concern about the relationship among error types, corrective feedback types, and learner uptake. In order to develop the positive washback of these ISLA studies, more explicit effort to address pedagogical implications with inservice and preservice teachers such as in Vásquez and Harvey (2010) seems necessary.

Second, despite an expansion of interaction literature, the linguistic targets in these studies are still mostly grammar, and target language is predominantly English. More replication studies with a variety of linguistic targets and in different languages would provide more insights and increase the validity of the findings. Furthermore, considering an increasing amount of research on multilingual speakers who possess multiple linguistic resources, it is pertinent to expand the scope of participants in interaction research by including multilingual speakers' learning of additional languages.

Recognizing the growing role that SCMC and other education technology is taking in the classroom, interaction research has begun to emerge in these instructional contexts.

Although increasing number of FTF interaction studies were conducted in intact classes as a part of their regular curricula (e.g., Kim, 2012), more SCMC studies that are conducted as a part of regular online courses are warranted to increase the ecological validity of ISLA research. This includes video- and audio-based SCMC. Furthermore, many interactional concepts in FTF mode introduced in this chapter (e.g., priming, task repetition, task modeling) have not been investigated in SCMC contexts. Future SCMC research is warranted to expand its research domain by exploring such topics.

The concept of interactive alignment (e.g., priming) has not been explored extensively in the SLA literature. Many L1 studies have shown the occurrence of interactive alignment, and more L2 studies focusing on this topic are needed. In particular, the delayed learning effects of priming during interaction warrant further investigation. With regards to priming effects between the two learners, much L1 priming research suggests that priming is based on implicit learning mechanism; however, L2 researchers have yet to address this question empirically. Future studies would need to explore whether it is an implicit language learning behavior as in L2 contexts L2 learners might explicitly try to copy other interlocutors' utterances to pursue successful conversations (i.e., explicit interactive alignment strategy). In terms of the literature on L2 priming tasks, the role of priming for application to classroom contexts is still in its infancy. Additional classroom-based priming research is needed to identify the most effective ways to design, sequence, and implement such collaborative tasks.

Finally, as previous interaction studies have suggested that interactional features are associated with language development, our ultimate goal is to train learners and teachers so that they can be good interaction participants in instructional contexts. For instance, Sato and Lyster (2012) trained learners to become efficient feedback providers during learner–learner interaction. ISLA research focusing on comparing different ways to train learners to become autonomous interaction participants is certainly a needed next step.

References

- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 29–51). Oxford: Oxford University Press.
- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 8, 543–574.
- Azkarai, A., & García Mayo, M. (in press). Task repetition effects on L1 use in EFL child task-based interaction. *Language Teaching Research*.
- Baralt, M., & Gurzynski-Weiss, L. (2011). Comparing learners' state anxiety during task-based interaction in computer-mediated and face-to-face communication. *Language Teaching Research*, 15, 201–229.
- Baralt, M., & Gurzynski-Weiss, L., & Kim, Y. (2016). Engagement with the language: How examining learners' affective and social engagement explains successful learner-generated attention to form. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 209–240). Amsterdam: John Benjamins.
- Brown, D. (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: A meta-analysis. *Language Teaching Research*, 20, 436–458.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Harlow: Longman.
- DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94–112). Routledge: New York.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281–318.
- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44, 449–491.
- Ferguson, C. A. (1971). Absence of copula and the notion of simplicity: A study of normal speech, baby talk, foreigner talk and pidgins. In D. Hymes (Ed.), *Pidginization and creolization of languages* (pp. 141–150). New York: Cambridge University Press.
- Fernández-García, M., & Martínez Arbelaiz, A. (2003). Learners' interactions: A comparison of oral and computer-assisted written conversations. *ReCALL*, 15, 113–136.
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3, 215–247.
- Fujii, A., Ziegler, N., & Mackey, A. (2016). Peer interaction and metacognitive instruction in the EFL classroom. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 63–89). Amsterdam: John Benjamins.
- Gass, S., & Mackey, A. (2007). Input, interaction and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175–200). Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S., & Mackey, A. (2015). Interaction approaches. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 180–206). New York: Routledge.
- Gass, S., & Varonis, E. (1994). Input, interaction and second language production. *Studies in Second Language Acquisition*, 16, 283–302.
- Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34, 445–474.
- Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in face-to-face and computer-mediated modes. *Studies in Second Language Acquisition*, 36, 1–37.
- Gurzynski-Weiss, L., & Baralt, M. (2015). Does type of modified output correspond to learner noticing of feedback? A closer look in face-to-face and computer-mediated task-based interaction. *Applied Psycholinguistics*, 36, 1393–1420.
- Hatch, E. (1978). Acquisition of syntax in a second language. In J. Richards (Ed.), *Understanding second and foreign language learning: Issues and approaches* (pp. 34–70). Rowley, MA: Newbury House.
- Iwasaki, J., & Oliver, R. (2003). Chatline interaction and negative feedback. *Australian Review of Applied Linguistics*, 17, 60–73.
- Jeon, K. (2007). Interaction-driven L2 learning: Characterizing linguistic development. In A. Mackey (Ed.), *Conversation interaction in second language acquisition: A collection of empirical studies* (pp. 379–403). Oxford: Oxford University Press.
- Junqueira, L., & Kim, Y. (2013). Exploring the relationship between training, beliefs and teachers' corrective feedback practices: A case study of a novice and an experienced ESL teacher. *Canadian Modern Language Review*, 61, 181–206.
- Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–131). Philadelphia, PA: John Benjamins.
- Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System*, 37, 254–268.
- Kim, Y. (2012). Task complexity, learning opportunities and Korean EFL learners' question development. *Studies in Second Language Acquisition*, 34, 627–658.
- Kim, Y. (2013a). Effects of pre task modelling on attention to form and question development. *TESOL Quarterly*, 47, 8–35.
- Kim, Y. (2013b). Promoting attention to form through task repetition in a Korean EFL context. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational settings* (pp. 3–24). Amsterdam: John Benjamins.

YouJin Kim

- Kim, Y. (2015). The role of tasks as vehicles for learning in classroom interaction. In N. Markee (Ed.), *Handbook of classroom discourse and interaction* (pp. 163–181). Malden, MA: Wiley-Blackwell.
- Kim, Y., & Taguchi, N. (2015). Promoting task-based pragmatics instruction in EFL classroom contexts: The role of task complexity. *Modern Language Journal*, 99, 656–677.
- Krashen, S. D. (1980). The input hypothesis. In J. Atlantis (Ed.), *Current issues in bilingual education* (pp. 168–180). Washington, DC: Georgetown University Press.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.
- Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology*, 10, 102–120.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal*, 97(3), 634–654.
- Loewen, S. (2012). The role of feedback. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 24–40). New York: Routledge.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Loewen, S., & Nabei, T. (2007). The effect of oral corrective feedback on implicit and explicit L2 knowledge. In A. Mackey (Ed.), *Conversational interaction and second language acquisition: A series of empirical studies* (pp. 361–378). Oxford: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in the adult English L2 classroom: Characteristics, explicitness, and effectiveness. *Modern Language Journal*, 90, 536–556.
- Loewen, S., & Reissner, S. (2009). A comparison of incidental focus on form in the second language classroom and chatroom. *Computer Assisted Language Learning*, 22, 101–111.
- Loewen, S., & Wolff, D. (2016). Peer interaction in F2F and CMC contexts. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 163–184). Amsterdam: John Benjamins.
- Long, M. H. (1981). Input, interaction, and second language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition* (pp. 250–278). New York: Annals of the New York Academy of Sciences.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269–300.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37–66.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265–302.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46, 1–40.
- Mackey, A. (1999). Input, interaction, and second language development: An empirical study of input, interaction, and second language development. *Studies in Second Language Acquisition*, 21, 557–588.
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27, 405–430.
- Mackey, A. (2012). *Input, interaction and corrective feedback in L2 classrooms*. Oxford: Oxford University Press.
- Mackey, A., Abbuhl, R., & Gass, S. (2013). Interactionist approach. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 7–24). New York: Routledge.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60, 501–533.

- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in SLA: A collection of empirical studies* (pp. 408–452). New York: Oxford University Press.
- Mackey, A., Philp, J., Fujii, A., Egi, T., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–208). Philadelphia, PA: John Benjamins.
- Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning*, 62, 704–740.
- McDonough, K. (2004). Learner–learner interaction during pair and small group activities in a Thai EFL context. *System*, 32, 207–224.
- McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition*, 27, 79–103.
- McDonough, K., & Chaikitmongkol, W. (2010). Collaborative syntactic priming activities and EFL learners' production of *Wh*- questions. *Canadian Modern Language Review*, 66, 817–842.
- McDonough, K., & Mackey, A. (Eds.). (2013). *Second language interaction in diverse educational contexts*. Amsterdam: John Benjamins.
- McDonough, K., Trofimovich, P., & Neumann, H. (2015). Eliciting production of L2 target structures through collaborative priming activities. *Canadian Modern Language Review*, 71, 75–95.
- McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. New York: Routledge.
- McDonough, K., & Trofimovich, P. (2012). Psycholinguistic methodologies for comprehension and production. In A. Mackey & S. Gass (Eds.), *A guide to research methods in second language acquisition* (pp. 117–138). London: Blackwell.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83–108.
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59, 411–452.
- Nassaji, H. (2016). Anniversary article interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, 20, 535–562.
- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17, 164–187.
- Nuevo, A. M. (2006). *Task complexity and interaction: L2 learning opportunities and development* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses database (UMI No. 3247335).
- O'Rourke, B. (2005). Form-focused interaction in online tandem learning. *CALICO Journal*, 22(3), 433–466.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109–148.
- Ortega, L. (2015). Second language learning explained? SLA across 10 contemporary theories. In B. Van Patten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 245–272). New York: Routledge.
- Park, S. (2010). The influence of pretask instructions and pretask planning on focus on form during Korean EFL task-based interaction. *Language Teaching Research*, 14, 9–26.
- Patanasorn, C. (2010). *Effect of procedural content and task repetition on accuracy and fluency in an EFL context* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff.
- Payne, J. S., & Ross, B. (2005). Synchronous CMC, working memory, and L2 oral proficiency development. *Language Learning & Technology*, 9, 34–54.
- Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, 20, 7–32.
- Pellettier, J. (2000). Negotiation in cyberspace: The role of chatting in the development of grammatical competence. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 59–86). Cambridge: Cambridge University Press.

- Philp, J., Adams, R., & Iwashita, N. (2014). *Peer interaction and second language learning*. New York: Routledge.
- Philp, J., Oliver, R., & Mackey, A. (2006). The impact of planning time on children's task-based interactions. *System*, 34, 547–565.
- Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21, 737–758.
- Plonsky, L., & Gass, S. (2011). Quantitative research methods, study quality, and outcomes: The case of interaction research. *Language Learning*, 61, 325–366.
- Plonsky, L., & Kim, Y. (2016). Task-based learner production: A substantive and methodological review. *Annual Review of Applied Linguistics*, 36, 73–97.
- Redmon, R. J., & Burger, M. (2004). Web CT discussion forums: Asynchronous group reflection of the student teaching experience. *Curriculum and Teaching Dialogue*, 6, 157–166.
- Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *Modern Language Journal*, 95, 162–181.
- Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge: Cambridge University Press.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–162). Amsterdam: John Benjamins.
- Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111–141). Amsterdam: John Benjamins.
- Sato, M., & Ballinger, S. (2016). Understanding peer interaction: Research synthesis and directions. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 1–30). Amsterdam: John Benjamins.
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34, 591–626.
- Sauro, S. (2011). SCMC for SLA: A research synthesis. *CALICO Journal*, 28, 1–23.
- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 43–57). New York: Oxford University Press.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8, 263–300.
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58, 835–874.
- Shekary, M., & Tahririan, M. H. (2006). Negotiation of meaning and noticing in text-based online chat. *Modern Language Journal*, 90, 557–573.
- Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 193–216). Amsterdam: John Benjamins.
- Smith, B. (2003). Computer-mediated negotiated interaction: An expanded model. *Modern Language Journal*, 87, 38–57.
- Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly*, 39, 33–58.
- Stockwell, G. (2010). Effects of multimodality in computer-mediated communication in tasks. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 83–104). London: Continuum International.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82, 320–337.
- Trofimovich, P. (2016). Interactive alignment: A teaching-friendly view of second language pronunciation learning. *Language Teaching*, 49, 411–422.
- Trofimovich, P., & McDonough, K. (Eds.). (2011). *Applying priming methods to L2 learning, teaching and research: Insights from psycholinguistics*. Amsterdam: John Benjamins.
- Trofimovich, P., McDonough, K., & Foote, J.A. (2014). Interactive alignment of multisyllabic stress patterns in a second language classroom. *TESOL Quarterly*, 48, 815–832.
- Trofimovich, P., McDonough, K., & Neumann, H. (2013). Using collaborative tasks to elicit auditory and structural priming. *TESOL Quarterly*, 47, 177–186.
- Truong, A. T., & Storch, N. (2007). Investigating group planning in preparation for oral presentation in an EFL class in Vietnam. *Regional Language Centre Journal*, 38, 104–124.
- Vásquez, C., & Harvey, J. (2010). Raising teachers' awareness about corrective feedback through research replication. *Language Teaching Research*, 14, 421–443.
- Williams, J. (1999). Learner-generated attention to form. *Language Learning*, 49, 583–625.
- Yuksel, D., & Inan, B. (2014). The effects of communication mode on negotiation of meaning and its noticing. *ReCALL*, 26, 333–354.
- Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition*, 38, 553–586.
- Zyzik, E., & Polio, C. (2008). Incidental focus on form in Spanish literature courses. *Modern Language Journal*, 92, 50–73.

Concept-Based Language Instruction

James P. Lantolf and Xian Zhang

Background

Sociocultural theory (SCT) has at times been thought not to have direct implications for instructed L2 development. Loewen (2015, p. 9), for example, pointed out that according to Ortega (2007), given its emphasis on joint assistance between learners, peers, and teachers, sociocultural theory assigns instruction a “complementary” rather than a core role in promoting language development. In that same chapter, Ortega grouped SCT with theories such as Associative-Cognitive CREED (Construction-based, Rational, Exemplar-driven, Emergent, and Dialectic), because in her view both theories consider instruction to be beneficial but neither theory proposes a specific instructional design (Ortega, 2007). We suspect that the source of what was a mischaracterization of SCT with regard to language pedagogy most likely is to be found in the fact that at the time most of the SCT-informed L2 research had used the theory as a lens through which interaction inside and outside of classrooms was analyzed. In fact, Lantolf and Thorne (2007) in the same volume did not foreground the importance of intentionally organized instruction for language development. Shortly before the publication of our chapter, Lantolf and Thorne (2006) included an extensive discussion of Negueruela’s (2003) dissertation, which signaled the beginning of the shift from using the theory as a lens to using the principles of SCT to systematically design language instruction (see Lantolf & Beckett, 2009 for discussion of the shift in orientation). Indeed, the majority of chapters in Lantolf and Poehner (2008) were dedicated to reports on the findings of instructional studies guided by SCT principles that were organized under two general headings: Dynamic Assessment (DA) and Concept-Based Instruction (CBI). Subsequently a series of studies has been carried out under the rubric of DA or CBI. Most, although not all, have been summarized and discussed in Lantolf and Poehner (2014), which argued for a praxis-based (unity of theory and practice) approach to classroom-based language development. Ortega (2015, p. 264), in the revised version of her earlier chapter, clearly recognized the transformation that had taken place as she grouped SCT-L2 with theories in which instruction seeks to “optimize” learning and “may be even necessary when the goal is truly advanced levels of proficiency.” She also appropriately acknowledged that SCT-informed pedagogy adheres to a particular instructional

design—Systemic-Theoretical Instruction, also known as Concept-Based Instruction (CBI), the topic of the present chapter.

Principles of Concept-Based Instruction

Vygotsky's approach to education in the new society that was under construction in the early years following the Russian Revolution, referred to as *developmental education*, has a very different meaning from its use in North America, where it typically refers to pedagogical interventions aimed at secondary or postsecondary students considered to be at risk for successfully completing a course of study in literacy, math, or science. For Vygotsky (1987) education refers to what he called the "artificial" development of the person through the intentional and systematic organization of conceptual knowledge that is optimally provided in formal education. In other words, for Vygotsky education is not an activity limited to the acquisition of new knowledge (i.e., learning); it is instead the activity that promotes a unique type of development generally unavailable in everyday life. Education promotes development by providing students with access to the type of scientific (also referred to as theoretical, or academic) concepts, which provide understanding of the object of study, whether it be mathematics, biology, physics, chemistry, history, art, or language that is deeper than what our everyday understanding may be. By and large, the latter type of knowledge is based on what cultures glean from direct observation and experience of the world through our senses. For example, our vision tells us that the sun rises in the east, moves across the sky, and sets in the west. Indeed, our language, as illustrated in the previous sentence, supports such a perspective. Science, as a consequence of a special type of rigorous analysis of the solar system, reveals a different understanding of planetary movement. Similarly, in our daily life, tomatoes, squash, eggplant, and cucumbers are classified as vegetables and we can expect to purchase these items in a vegetable market. Botany tells a different story, however. All of these objects are fruit and share with apples, oranges, grapes, and peaches the fact that they are seed pods surrounded by pulpy flesh (i.e., ovaries). Consider another distinction between everyday and scientific knowledge. Although we have most likely observed and probably have experienced the movement of people in a crowd and the flow of water through narrow openings such as a gorge or a simple garden hose, we are no doubt unaware that both types of movement are related and are in fact explained by the same principle of fluid dynamics known as Bernoulli's law, which also explains the lift that raises airplanes off the ground (Kinard & Kozulin, 2008).

In the case of language, we certainly acquire everyday understandings of our primary communicative system that are often saturated with ideology such as the mistaken assumptions that a community's language will become corrupted if foreign words are adopted by its speakers, or that bilinguals are not only linguistically but also cognitively defective. It is also the case, with regard to second and foreign language instruction, that much of the explicit knowledge that learners are provided in formal instruction is often incomplete, misleading, or closely linked to specific contexts of use and therefore leads to problems when learners attempt to generalize this knowledge to other contexts. Negueruela (2003), for example, documented that students with previous study of Spanish in secondary school had internalized an understanding of Spanish verbal aspect (i.e., preterit and imperfect morphology in past tense) that was either wrong or too restrictive. Moreover, when they were provided with more accurate and complete conceptual knowledge of the meaning of aspect as temporal perspective, the students had difficulty coping with the new knowledge as it conflicted with their previous learning. According to Miller (2011), unless education is able to replace

James P. Lantolf and Xian Zhang

the pre-understanding that students bring to the process (whether this originates from their everyday experiences or from previous instruction) with new scientifically grounded conceptual knowledge, development is likely to be hindered or blocked altogether.

This brings us to another key principle of developmental education—the importance of instruction for development. From at least the middle of the 19th century, beginning with the writings of Herbert Spencer, educators assumed that successful instruction depended crucially on student readiness to learn (Egan, 2002). Consequently, not only was instruction designed to tap into the abilities that students had already developed through their interactions with the everyday world, it attempted to import into schooling those very processes that children use when learning outside of school. This orientation to learning provided justification for what is often referred to as discovery or inquiry-based education, where students are encouraged to accumulate knowledge through exploration of particular aspects of the object of study (Karpov, 2014). Moreover, discovery learning propagates the assumption that children develop according to a built-in set of abilities that emerges according to a biologically specified time table that is unaffected by teaching (Egan, 2002). This commitment to the so-called *natural child* has had its influence on various theories of SLA and L2 education, including Krashen's (2000) monitor model and natural approach and Pienemann's (1998) processability theory and teachability hypothesis (Pienemann, 1989).

Vygotsky (1987) recognized that biology is of course implicated in the formation of human consciousness. A human brain and body are necessary components for human development. Indeed, he argued, humans share many mental abilities, such as memory, attention, reflexes, and biological urges with higher primates; however, these abilities are not what make us human. It is rather the dialectical interweaving of these capacities with culturally created forms of mediation that give rise to specifically human forms of consciousness, whereby humans develop the ability to control their mental processes as a result of participating in culturally mediated activities. In other words, humans do not just remember, pay attention to, and perceive things in the world: they do so intentionally and with specific goals in mind and in accordance with the social relations they participate in as organized by the institutions of their culture, including family life, religious organizations, work, political and economic life, and education. Hence, humans do not relate directly to the world, as is the case with animals; they indirectly relate to, and act upon, the world through the appropriation of specific forms of cultural mediation. Thus, a central principle of SCT is that human thinking does not emerge as a consequence of biological maturation of the brain, but develops as a result of the appropriation and internalization of sociocultural forms of mediation.

Key Concept

Mediation: social relations and cultural artifacts such as language that humans appropriate from others to organize and control their mental processes.

As the forms of mediation that people have access to change, the possibilities for how they mediate their mental processes also change. One of the most impactful differences in mediation is that between everyday life and formal education. In school two forms of mediation, instruction (a specific type of social relation) and conceptual knowledge (the result of rigorous scientific analysis), are intentionally and systematically organized in ways that are

quite different from what occurs in everyday life. Education, according to Vygotsky, entails a special form of developmental activity that individuals do not normally have access to in the everyday world. In effect, his argument is that rather than waiting for students to become developmentally ready to learn, educational activity itself is the process through which development occurs. The term that Vygotsky used to describe what happens in school is *obuchenie*, the process in which instruction, dialectically intertwined with learning, leads development (Cole, 2009). Instruction provides access to high quality conceptual knowledge and is sensitive to the learners' *zone of proximal development*.

Key Concept

Zone of proximal development: a projection of an individual's (or group's) future development based on what they are able to do alone, actual development, and what they are able to achieve with appropriate mediation. It is the process through which internalization occurs.

Although Vygotsky laid the foundation for developmental education and concept-based instruction, he did not provide concrete recommendations for how it might be effectively implemented. This task was taken up by two of his students, Piotr Gal'perin and Vasily Davydov, each of whom had slightly different approaches to the process. For purposes of the present chapter, we will focus on Gal'perin's proposal, which has thus far been favored by L2 researchers (see Davydov, 2004 for a discussion of his pedagogical model).

Gal'perin's Approach to Concept-Based Instruction

Gal'perin and members of his research group carried out nearly 800 pedagogical studies (summarized in Talyzina, 1981) in a wide array of school subjects, including math, geometry, physics, and language, out of which emerged five theoretically informed and empirically supported instructional recommendations. Gal'perin and colleagues referred to their approach to developmental education as System-Theoretical Instruction (Haenen, 1996), although as we mentioned earlier, we will follow the current practice of referring to the approach as Concept-Based Instruction (CBI). The goal of CBI is to provoke development through effective presentation of high-quality conceptual knowledge connected to practical activity whereby the students not only internalize the conceptual knowledge but also come to understand how they can deploy the knowledge to meet their own goals in any of the subject areas taught in school, including language. In the following paragraphs we briefly describe each of the five recommendations that constitute Gal'perin's approach to developmental education—CBI.

The initial phase is referred to as the *orienting* phase of an action, which determines the overall quality of the action (Gal'perin, 1969, 1989a, 1989b, 1992). Orientation entails an ability to plan one's actions symbolically prior to objectifying them in the material world. The orienting phase of a physical or symbolic action necessitates an intention to realize a particular goal, knowledge of how to achieve the goal, and access to resources needed to realize the appropriate goal-directed action in the material world (Gal'perin, 1969, 1979). The action may range from architects designing buildings, families rearranging furniture, or a speaker/writer constructing a message. The knowledge necessary for the orienting phase

James P. Lantolf and Xian Zhang

of an action of course can be of the everyday, or of the scientific/theoretical, variety. Following Vygotsky's principles, Gal'perin proposed that to maximize the effectiveness of developmental education, scientific knowledge must be provided as "a meaningful whole" and avoid a purely verbal format, which often encourages rote memorization without genuine understanding (Haenen, 2001, p. 162). Given Vygotsky's stance regarding education leading, rather than following, development, the knowledge must be future oriented; that is, it must have utilitarian value so that students can use it to regulate their future behavior (physical or symbolic) in a virtually unlimited array of activities. In this sense, knowledge that is future oriented awakens the students' ZPD (Haenen, 2001). In other words, the new knowledge creates dissonance with what learners already know (i.e., actual level of development) and compels them to find ways of reconciling the discrepancy. They may, of course, ignore the new knowledge altogether, in which case development will not occur; or they may, under the mediation of a teacher, appropriate the knowledge in a way that either rejects their previous understanding (i.e., old knowledge), or they may integrate the knowledge with their previous understanding. Either of these latter options is considered to constitute development in the ZPD.

With regard to language development, internalization implies control of conceptual knowledge of the target language, especially concerning the meaning-making possibilities afforded by language. Thus, CBI strongly encourages reliance on theories of language that privilege meaning rather than structure. Consequently, Lantolf (2011) and Lantolf and Poehner (2014) argued in favor of theories that privilege meaning over form, including cognitive linguistics, systemic functional linguistics, and usage-based approaches. Cognitive linguistics (see Tyler, 2012) is particularly attractive because its theoretical explanations are generally formulated in graphic form that often can be easily modified to generate viable SCOBAs (Schemas for the Orienting Basis of Action). Nevertheless, other meaning-based theories of language are useful sources of conceptual knowledge.

Gal'perin's proposal for formatting the presentation of theoretical knowledge was captured in his concept of SCOBAs, or Schema for the Orienting Basis of Action (Gal'perin, 1989b), which makes up the second phase of his model. A SCOBAs is a visual, and if possible, material, holistic explanation of scientific knowledge that enhances understanding, is memorable, but at the same time mitigates the likelihood of rote memorization, and enables learners to use the knowledge in practical activity. This is not to say that linguistically based explanations of concepts are not part of the educational process, but these are linked to, and eventually replaced by, SCOBAs. A bit later in the chapter we will illustrate SCOBAs in conjunction with our discussion of a study on teaching Chinese pragmatic word order.

Key Concept

SCOBAs—Schema for the Orienting Basis of Action: a holistic visual or material representation of scientific knowledge that enhances student understanding and at the same time is memorable and functional.

We pointed out earlier that SCOBAs need to be memorable; that is, students must eventually be able to retain the knowledge represented in a SCOBAs in the physical absence of the SCOBAs. Students cannot be expected to physically transport the SCOBAs from one

communicative activity to another, although they are quite likely to need do this at the beginning of the developmental process. Said in another way, at the outset of instruction, a SCOBAs is an external form of mediation that students use to regulate their communicative behavior; over time, the knowledge it represents must be internalized in order to efficiently generalize to future communicative activities in and out of the classroom. Valsiner (1997) defined internalization as follows:

Internalization is a negotiated process of development that is co-constructed through constant forward-oriented construction of signs that bring over from the extrapersonal (social) world of the person to the intrapersonal subjective world semiotically encoded experiences, which, as personal sense systems, guide the person's process of further reorganization of the person-environment relationship.

p. 246

Although co-construction, mentioned in the definition, is usually construed as participation of other individuals (e.g., parents, siblings, teachers, peers), it can also imply involvement of other extrapersonal forms of mediation, such as SCOBAs, which, because they are constructed by someone, also qualify as social forms of mediation (see Vygotsky, 1978).

The third phase of the model connects SCOBAs to practical activity that promotes their internalization and the conceptual knowledge they depict. In the case of language, practical activities should be designed to enable learners to use the relevant concepts to carry out particular communicative goals. These activities may involve tasks, scenarios (Di Pietro, 1987), drama, etc., that incorporate experience with different language modalities and that cover a range of language features including grammar, discourse, pragmatics, and figurative language.

The fourth phase reflects Vygotsky's (1987) view of language as a psychological tool whereby the relevance of overt and covert verbalization for development is integrated into the model. Swain (2006, p. 96) referred to this mediational function of language as *languageing* and described it as "producing language in an attempt to understand—to problem solve—to make [personal] meaning." Gal'perin supported two different languageing formats: (1) *communicated thinking*, whereby learners explain their understanding and use of concepts to someone else (e.g., a teacher or a peer)—"I"/"You" interaction; and (2) *dialogic thinking*, whereby the explanations are directed at the self—"I"/"Me" private speech (Haenen, 2001). Thus, the process of internalization moves from social to psychological activity (Vygotsky, 1987).

The language phase leads to the fifth and final phase of internalization where the concept can now be deployed with relative fluency in a variety of communicative (spoken and written) activities. This phase is also known as the inner speech phase because the planning or orienting process is carried out through inner speech.

Current Issues

One issue that has been raised regarding CBI is its assumption that educators have the requisite subject-matter knowledge to be able to explain concepts appropriately and then visualize/materialize these as functional SCOBAs. In the case of language, as argued in Lantolf and Poehner (2014), this assumption may well necessitate rethinking teacher education programs in order to provide more extensive opportunities for teachers to develop the kind of conceptual knowledge needed to implement CBI. Lantolf and Poehner (2014) also suggested

that conceptual knowledge is equally if not more important for effective instruction than is teacher proficiency in the language. This is not to argue against teacher proficiency, but it is to argue against the view that language proficiency on the part of the teacher is in itself sufficient to enable a teacher to function as an effective language instructor. It seems rather evident that native-speaker ability without in-depth explicit knowledge of how the various features of a language function is very likely to result in rather naïve and unsystematic explanations of how the features operate. While native speakers, per se, may be able to indicate that different uses of their language may or may not be acceptable and appropriate, they are unable to provide the necessary conceptual knowledge that is needed to help classroom learners develop communicative ability in the language without explicit language-focused preparation. As Lantolf and Poehner (2014) point out, the majority of teacher education programs fail to provide this type of preparation.

A second issue is the fact that Gal'perin seemed to assume that conceptual knowledge of an educational topic must necessarily originate from visualized/materialized conceptual knowledge linked to concrete practical activity. Gal'perin also acknowledged that student development can indeed result from reading texts on one's own, imitating others and listening to explanations, but he nevertheless cautioned that under these conditions there is a real danger that students will have difficulty segregating essential from nonessential features of the object of study (Gal'perin, 1989b), as frequently occurs in rules-of-thumb pedagogy (see Negueruela, 2003).

Empirical Evidence

In this section we will first briefly review some of the recent studies carried out using CBI for L2 classroom development. We will then focus on one study that addressed pragmatic word order in L2 Chinese (see Zhang, 2014; Zhang & Lantolf, 2015). We will review the study but will then consider an aspect of the study that to date has not been discussed in the published literature—the influence of CBI on working memory.

It is difficult to provide an exact count of the CBI studies on language that have been carried out since Gal'perin formulated his educational model. Many of the early studies were carried out in the Soviet Union and have not been well documented in Western research literature. However, in the 1970s some studies were conducted in Western Europe involving languages such as German, French, and Russian. Most of these were fairly short-term studies lasting no more than a few hours (for details see Lantolf & Poehner, 2014). However, beginning with Negueruela's (2003) semester-long study of Spanish aspect, mood, and modality, an increasing number of extensive classroom L2 projects, primarily as doctoral dissertations, have been completed. These projects have focused not only on grammar, but also on pragmatics, and figurative language, as well as reading and writing. The languages that have been the object of instruction include English, Spanish, Chinese, and French. Space does not permit a full in-depth review of the CBI L2 research; instead, we will provide a representative sample with one study that focuses on grammar, one that deals with pragmatics, and one that addresses figurative language. The interested reader can find more details in Lantolf and Poehner (2008, 2014) and van Compernelle (2014).

The first study, carried out by Lee (2012), addressed the grammar of English phrasal verbs composed with the particles *over*, *out*, and *up*. The study was conducted in an intact university intermediate level ESL class with an enrollment of 23 students, whose L1s were either Chinese, Korean, or Thai. Lee was the instructor for the course, which followed the mandated syllabus, with additional CBI for phrasal verbs carried out over a 3-week period

near the end of the 15-week semester. The project began from the assumption that phrasal verbs are grounded in semantic principles that can provide a motivated explanation for the seemingly random combination of verbs and particles, which avoids the often relied on pedagogical practice of learners memorizing lists of verb + particle collocations. Given that cognitive linguistics stresses meaning over form, it was felt that this theory of language analysis would be able to provide the conceptual foundation for a pedagogically viable explanation that illustrated the relationships between the literal and metaphorical extensions of verb + particle couplings.

Here we will briefly consider only one of the verbal particles, *out*. Basing her analysis of the particle on the work of Morgan (1997), Lee noted that *out* presupposes the existence of a container, either literally or metaphorically, and therefore abstract domains of use can be conceptualized as containers, which provides a systematic explanation for what seemingly appears as an arbitrary random use of the verbal + particle combination. Many abstract conceptual domains can be conceptualized as containers, and this conceptualization provides a more systematic explanation for the seemingly random use of *out*. In an utterance such as (1), *I took the glass out of the cupboard*, the verb and the particle both retain their literal senses where the meaning is literal movement out of a literal container. In (2), *She fished out the ring*, the verb has a metaphorical meaning (she did not literally catch a fish), while the particle keeps its literal meaning of movement out of a container. In (3), *We picked out a name for the baby*, both verb and particle are metaphorical, because the name was not actually picked, as when someone selects a piece of candy from a box, and nothing literally moved “out” of a container, as would be the case for the candy. There are other meanings associated with *out*, which Lee included in her study, but a discussion of these would require much more additional space. The point to be made here is that conceptually, if one understands the basic literal meaning of the combination of a verb + *out*, one is likely able to determine its meaning in metaphorical extensions. Lee then developed a SCOPA to depict the possible meanings of *out* combined with verbs. In Figure 9.1 we illustrate one component of the meaning of *out*, specifically as it relates to selection from a larger group (e.g., fish out a ring).

Using a pretest/posttest procedure, Lee showed that not only did learners improve in their ability to correctly interpret verb + particle combinations, they also developed a greater sense of confidence in their interpretations. At the outset of instruction, the students indicated that their primary interpretive strategy was to guess at the meaning of phrasal verbs. Moreover, on the posttest the students were not only able to correctly determine the meaning of the items included on the pretest, they were also able to better understand new verb + particle combinations, including those constructed with two particles, *in* and *down*, that were not addressed during instruction. According to Lee (2012) the increase in learner interpretive confidence as well as in performance, including most especially with regard to the two new particles, is a clear indication that learners had internalized the conceptual knowledge regarding how meaning in phrasal verbs is constructed.

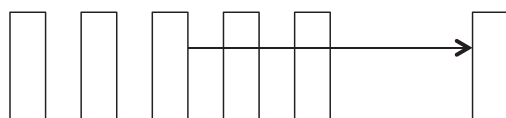


Figure 9.1 Partial SCOPA for particle “out,” meaning “selection”

The next study (van Compernelle, 2012, 2014) focused on the pragmatics of French second person address forms (*tu/vous*), negation (*ne . . . pas/. . . pas*), and first person plural pronouns (*on/nous*). In the interest of space, we will limit our discussion to the first of the three topics. The study was carried out in a one-on-one tutorial format over a 6-week period and involved eight fourth-semester university students of L2 French with a focus on speaking and reading proficiency. As van Compernelle (2012) pointed out, learners with classroom experience only by and large do not have much awareness of the subtleties regarding how to appropriately use pragmatic variants such as those included in the study.

The fundamental concept for van Compernelle's (2012) study was derived from Silverstein's (2003) notion of *orders of indexicality*, which refers to the nonlinguistic interactional meaning that communicators use to mark the various kinds of social factors that enter into communicative exchanges. Using three separate index cards, van Compernelle (2012) created an indexicality SCOPA in which first order indexicality was explained as language use "based on geographic location, formality of context, age of speaker, level of education, and social class" (p. 66); second order indexicality was explained as the conventions people use to sound local, formal, younger or older, more or less educated, more or less high class, or like any one relevant group of speakers; and third order indexicality entailed noticing and valuating how people use second order indexicals. Van Compernelle created five SCOPAs to visualize how the three orders of indexicality are manifested in communicative interaction with regard to the three pragmatic topics addressed in the study.

The SCOPA for *tu/vous* depicted interlocutors in three potential interdependent configurations. The first illustrated informality/formality through individuals dressed in, respectively, T-shirt and jeans, and suit and tie; the second depicted social closeness by situating both interlocutors within the same picture and social distance by situating each interlocutor in a separate picture; the third configuration, relative social status, indicated equality by situating the picture of each interlocutor at the same level and depicted their inequality by positioning each interlocutor on a different level (i.e., slightly higher than the other). Thus, for example, if interlocutors wished to index informality (T-shirt and jeans), lack of social distance, and equal social status they would opt to mark their relationship through mutual use of *tu*. If, on the other hand, they chose to index informality, social distance, and equal social status they would address each other as *vous*. As van Compernelle (2012) pointed out, the possible indexical combinations are more complex than what is described in most French textbooks, and more importantly, they are determined not by some hard-and-fast rule of thumb but are negotiated by interlocutors during actual face-to-face communication.

Following presentation of the concept of indexicality and the SCOPAs, van Compernelle (2012) engaged the students in a series of scenarios (see Di Pietro, 1987), which were a kind of unscripted mini-drama that revolved around a conflict (e.g., one interlocutor wishing to smoke an e-cigarette and the other not able to tolerate smoke of any kind). After each scenario was performed the students were asked to explain the basis of the choices they made within each of the three pragmatic options (address, negative, first person plural). Among other things, van Compernelle (2012) reported that the students had a greater degree of understanding of how indexicality functions in French and that this allowed them to be more flexible in making their pragmatic choices during actual performance.

The third study, carried out by Kim (2013), addressed the topic of ESL learner identification and interpretation of spoken sarcasm, a domain, which according to Kim, causes considerable difficulty for students, even at fairly advanced levels of proficiency. Sarcasm is a type of irony in which a speaker says one thing and means another with either a humorous

or insulting and critical intent. The problem is that there are differences in frequency of sarcasm use across speech communities, and not all speech communities use the trope to indicate humor, instead reserving it for insult and criticism. Additionally the cues, both verbal and nonverbal, that mark sarcasm vary across communities. The Anglo world makes heavy use of sarcasm in both its humorous and negative functions, and certainly to a much greater extent than do speakers of Korean. Kim's project involved nine speakers of L1 Korean who were advanced speakers of L2 English enrolled in various graduate programs at a large US research university. Instruction was carried out primarily in Korean and lasted for 16 weeks. Kim formulated eight SCOBAs that illustrated how to detect and appropriately interpret spoken sarcasm in English. The SCOBAs depicted such features as tone of voice, physical stance (e.g., head tilted to the side, hands on hips, facial expression), lexical indicators (e.g., yeah, yeah), and so forth that signaled sarcastic intent on the part of a speaker. She created a series of video clips from YouTube, including excerpts from American sitcoms, which illustrated that various ways of marking sarcasm and how to interpret the positive (i.e., humorous) or negative (i.e., insulting) intent of the speaker.

Kim used a pretest, posttest, and delayed posttest (administered 1 month after instruction) to assess the learners' ability to detect and appropriately interpret sarcasm. The tests comprised a series of video clips, some of which depicted sarcasm and some of which did not. She administered the tests to native speakers of English to determine the validity of the test. The posttests incorporated items that had not appeared on the pretest. All learners improved significantly from pretest to posttest and maintained their ability on the delayed posttest. Moreover, in post-instruction interviews, the students indicated that they felt more empowered when interacting with native speakers of English as a result of instruction, and equally important, the learners reported an enhanced understanding and sensitivity to sarcasm in Korean, their native language. One of the points Vygotsky (1987) made with regard to formal education is that the study of additional language can result in enhanced understanding of learners' native language.

CBI, Chinese Word Order, and the Teachability Hypothesis

In his dissertation on the pragmatics of Chinese word order and the Teachability Hypothesis, Zhang (2014) demonstrated that it was possible, through instruction, not only to mediate the development of learner ability to manipulate word order in Chinese but also to assess Pienemann's (1989) Teachability Hypothesis from the perspective of CBI. Briefly, the Teachability Hypothesis, based on Pienemann's (1998) Processability Theory, argues that instruction cannot interfere with the presumed natural developmental sequence that learners follow when acquiring specific features of an L2. Furthermore, if teaching is to be effective with regard to these features it should be aimed at the next stage in the processing hierarchy. Thus, if for a particular feature that is subject to processing constraints (e.g., question formation in English, word order in German) a learner is at stage 2, for example, instruction can prepare a learner to reach stage 3 but it cannot provoke the learner to skip to stage 4. This is so because stage 3 is considered to be the necessary prerequisite for processing ability at stage 4. The processing constraints are governed both by specific linguistic principles as specified in lexical functional grammar and by psycholinguistic factors proposed in Levelt's (1989) model of speech production (see Pienemann, 1998 for details). Some recent research has demonstrated, however, that instruction can promote development from one processing stage to the next even if it aims at stage $X + 2$. Nevertheless, the assumption is that learners still cannot skip an intervening stage (Bonilla, 2012).

James P. Lantolf and Xian Zhang

In Zhang's (2014) study, he designed an instructional program on the pragmatics of Chinese word order (explained later) that resulted in learner ability to skip a processing stage. Specifically, Zhang reported that learners entering instruction at stage 2 could develop the ability to process stage 4 features before stage 3 features and that it was also possible to mediate learners into the ability to process stage 3 and 4 features simultaneously. According to Pienemann's (1998) criterion for the emergence of processing ability at a given stage, learners are not expected to use features from the stage with a high degree of accuracy (e.g., 80% to 90%); rather they are expected to use the feature accurately in at least four or five different contexts where the feature normally occurs.

Zhang (2014) worked with four undergraduate university L1 English learners of Chinese. The feature of interest was variation in word order to indicate topic/comment information. Basic word order in Chinese is SVO, as in (1), where the subject is also the topic of the utterance:

- (1) *tā hē le kāfēi.*
He drank coffee.

In basic word order, when a temporal or locative adverb is used, the anticipated order is S Adv V O as indicated in (2):

- (2) *tā zǎo shàng hē le kāfēi.*
He morning drank coffee.

However, if a temporal or locative adverb functions as the topic, as for instance when answering the question *When did he drink coffee?*, the adverb appears in the initial position as in (3):

- (3) *zǎo shàng tā hē le kāfēi.*
morning he drank coffee.

If a speaker wishes to mark a sentential object as the topic of an utterance, as when responding to the question "What did he drink in the morning?" the object appears in the utterance initial position, as in (4):

- (4) *kāfēi tā zǎo shàng hē le.*
coffee he morning drank.

According to the processing hierarchy, S (Adv) VO order is a stage 2 ability, while Adv SVO is stage 3 and O S (ADV) V is a stage 4 ability. Stage 2 is canonical word order in Chinese; that is, it positions the subject in initial, or topic position. Adverbs, either of time or location, if they are relevant to what a speaker wishes to say, are positioned between S and V. According to Processability Theory, this stage requires less processing capacity than stage 3, whereby a speaker must situate a constituent in sentence initial position—the position where the S normally appears. Stage 4 requires even more capacity to process because it topicalizes a constituent, O, which must first be marked for case by the verb that governs it before it can be repositioned. Adverbs, in general, are easier to process than are verb arguments, such as O, because they are generally not governed by another syntactic category.

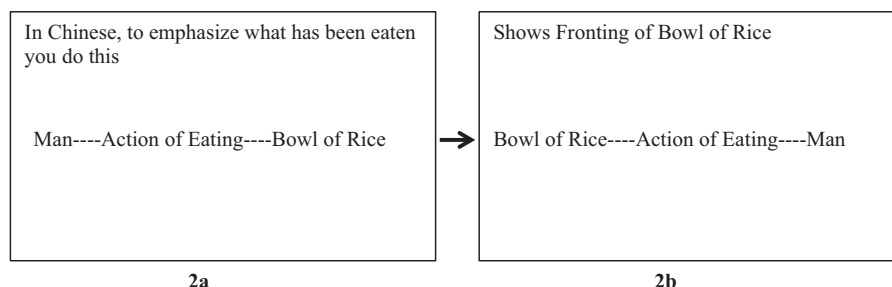


Figure 9.2 SCOPA for topicalization of object

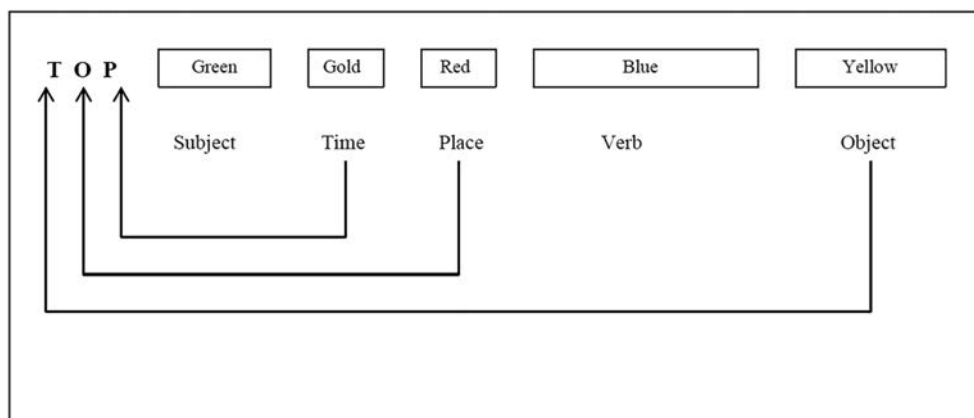


Figure 9.3 Material SCOPA illustrating topicalization options in Chinese

Zhang developed three different oral production instruments to assess emergence of the three relevant processing stages: elicited imitation, a question and answer interview (Q&A), and a cartoon narration. He created a series of SCOPAs that illustrated word-order variation in Chinese. Figure 9.2 is the SCOPA for positioning O in utterance initial position.

The SCOPA was presented in an animated PowerPoint (PPT) format in which the rice moved from third position in 1a to first position in 1b. A similar display was used to illustrate topicalization of locative and temporal adverbs.

In addition, Zhang also used Cuisenaire rods, associated with Silent Way pedagogy (Gattegno, 1963), to create the material SCOPA shown in Figure 9.3 to engage the learners in the activity of physically manipulating the topicalization options of Chinese. For convenience we show the colors orthographically.

The rods varied in color (matching the PPT displays) and in length as a way of indicating their grammatical status within an utterance. S and O are of the same length as they indicate different nominal arguments that co-occur with the verb in transitive constructions. The two adverbs are smaller than the other rods in order to indicate their optional status. The verb is the largest rod, which indicated that it remains *in situ* in Chinese utterances. The arrows indicate the constituents (i.e., object, temporal, locative adverbs) that can be topicalized. It was explained and demonstrated with SCOPAs such as depicted in Figure 9.2 that only one

James P. Lantolf and Xian Zhang

constituent may be topicalized in any utterance. It was further explained to students that S in canonical S (Adv) V O constructions takes on the topic function.

Following administration of the three pretests (elicited imitation, Q&A, and cartoon narration), which confirmed that all learners were at stage 2 (i.e., SVO), the students were given instruction on stage 4 OSV structures. The instruction involved explanation of the meaning of topicalization and how it is manifested in Chinese with regard to O only. The SCOA in Figure 9.2 was presented. The learners then carried out various practice activities, including sentence construction, gap filling, Q&A, translation, cartoon narration, and free talk in which they had the opportunity to discuss any topic of personal interest with the tutor, including comments on their roommates, issues that came up in other classes, and so forth. During the instructional phase learners also demonstrated their understanding and use of O topicalization not only by responding to the practice activities orally but also by physically manipulating the rods as in Figure 9.3. They also took part in languaging activities where they were asked to explain to the instructor in English their understanding of the concept of topicalization and why they used it in specific activities, such as cartoon narration and Q&A. The following week, learners were given the posttests for O topicalization followed by instruction on Adverb topicalization, stage 3 in the processing hierarchy. The same instructional procedure was followed.

On the first posttest for O topicalization none of the learners produced Adverb topicalization even though contexts where adverbs could have been topicalized were available. This result is counter to the findings of research such as Bonilla's (2012) in which instruction carried out at an $X + 2$ stage (in Zhang's study $X = \text{SVO}$ and $+ 2 = \text{OSV}$) was effective but only in moving learners to the next $X + 1$ stage in the processing hierarchy. Following Adv instruction (stage 3), the learners were given a posttest 1 week later to determine if they were able to process topicalization at both stages in the hierarchy. The results verified that indeed the performance of the four learners met Pienemann's processing criterion—using both stages in at least four different contexts (see Zhang & Lantolf, 2015). The same result was found for the delayed posttest administered 1 month following instruction. Consequently, Zhang's (2014) study demonstrated that CBI was an effective instructional approach, and as proposed by Vygotsky, that systematically organized instruction can result in developmental processes that are different from those that occur in the less systematically organized activity of everyday life when learners are not always offered the kind of mediation and support that developmental education provides.

Teaching Tips

- When explaining language features to students, focus on how the features provide options for conveying meaning rather than focusing on the form of the feature.
- Formulate visual means of representing a language concept for students rather than explaining the concept in words only. Visual representations are more easily remembered than are verbal representations.
- Provide opportunities for students to talk about their understanding of new language concepts with each other and with you in order for them to develop a deeper understanding of the concept.

CBI and Working Memory

To further document the effects of CBI on development, we would now like to consider another piece of evidence from Zhang's (2014) study—the impact of CBI on working memory (WM), considered by SLA researchers to be an important factor in successful learning outcomes (see Williams, 2012). WM “refers to the [mental] system or systems that are assumed to be necessary in order to keep things in mind while performing complex tasks such as reasoning, comprehension and learning” (Baddeley, 2010, p. 136). WM supports human thinking by serving as “an interface between perception, long-term memory and action” (Baddeley, 2003, p. 191). WM comprises four components that deal with real time cognitive processing: the central executive, the phonological loop, the visuospatial sketchpad (VSSP), and the episodic buffer (Baddeley, 2000). Rather than delve further into a discussion of the function of working memory, in the interest of space we refer the reader to Chapter 22 by Li where the author discusses WM and its relevance for L2 learning, especially with regard to the significant correlations that have been reported in the literature between grammatical ability and WM, and most importantly, for our purposes, at the beginning levels of proficiency.

We focus our attention on the performance of one learner who manifested a low level of performance on both the English and Chinese WM tasks administered by Zhang. On the English version of the phonological-loop task, the mean score for all learners participating in the Zhang's project was 13.3 ($SD = 2.6$) with a range of 8 to 16. On the Chinese version of the phonological-loop task the mean score was 7.7 ($SD = 3.3$) with a range of 2–14. In both tasks the maximum score possible was 21. The student of interest, whose pseudonym is Kris, produced the lowest score on both the English (8) and the Chinese (2) tasks. Moreover, on a Chinese vocabulary recognition task, Kris recognized 49 words, again the lowest score, while the mean recognition score for the group was 60. Given the findings reported in Li (this volume) with regard to WM and L2 proficiency, we would expect Kris not to perform well on Zhang's posttest tasks when compared to the students with higher WM scores. As it turned out, however, this was not the case.

Kris's performance on the posttests indicated that she had indeed developed the ability to process stage 4 OSV structures on a par with the other learners in the study. For example, on the delayed posttest Amy (pseudonym), whose WM scores were 13 for English and 9 for Chinese, produced a total of 26 OSV utterances out of 121 contexts of use, while Kris produced 33 OSV utterances out of a possible 113 contexts where the structure could have been used. Both learners far exceeded Processability Theory's criterion for processing ability.

The interesting question is whether or not learners, such as Kris, with a small phonological loop can compensate in some way for their apparent disadvantage when it comes to language learning. While it has been generally assumed that WM capacity is a stable cognitive trait, some research has nevertheless suggested that it may be possible for some people to devise strategies to compensate for a poor WM (see Gathercole, Tiffany, Briscoe, & Thorn, 2005), while other studies have documented the positive effects of training on the improvement of reduced WM capacity (see Holmes, Gathercole, & Dunning, 2009; Klingberg, Forssberg, & Westerberg, 2002). Besides training aimed at enhancing the capacity of the phonological loop, another solution to WM problems may be to recruit other components of the WM system to work in conjunction with the phonological loop. Instruction that taps into a learner's VSSP for instance, may be an effective means for overcoming the disadvantage of a small phonological loop. For learners with a small WM, maintaining the information necessary to

James P. Lantolf and Xian Zhang

produce an utterance, including lexical, phonological, and grammatical knowledge, could easily overload the phonological loop system.

The VSSP, not discussed in much detail in Li's chapter, may be a possible system that allows an individual to compensate for a small phonological loop. The VSSP is specialized for maintaining visual and spatial information in short-term memory and is therefore associated with nonverbal intelligence, but it might also facilitate the acquisition of semantic knowledge relating to the appearance and use of objects such as machinery as well as spatial orientation and geographical knowledge (Baddeley, 2003). We believe that the SCOBAs utilized in Zhang's study, and in particular the materialization represented through the Cuisenaire rods, were influential in extending Kris's WM. The SCOBAs could well have recruited Kris's VSSP, which, as mentioned, promotes the conversion of visual semantic information into long-term memory. This type of long-term memory store in turn could have assisted Kris in generating OSV utterances.

Kris mentioned that whenever she encountered problems producing an appropriate utterance she would visualize the rods. On occasion she even produced co-expressive gestures as if she were manipulating the rods (see Figure 9.4), even though the rods were not physically present during the posttests. When describing a scene in the cartoon narration posttest in which the cartoon character, Jerry the mouse, drank a bowl of soup, Kris simultaneously produced the Chinese utterance *Soup Jerry drank* (OSV) while her right hand was positioned on a table, as seen in Figure 9.4. Her right index finger and thumb formed a shape that resembled a rod (presumably the yellow rod in Figure 9.3). She maintained this shape while moving her hand from right to left in order to reposition the Object into utterance-initial



Figure 9.4 Kris's gesture while producing an OSV utterance

position. The co-expressive gesture provided a window into Kris's cognitive processing (see McNeil, 2005). Because visual information is stored in the VSSP, we hypothesize that Kris recruited this component of WM to help her deal with the complexity created by her impoverished phonological loop. Recruiting the VSSP allowed Kris to distribute information between the phonological loop and the VSSP. The phonological loop stored the lexicon that Kris needed to produce the sentence. The VSSP stored the word order information as shown by her co-expressive gesture. The rods, together with the SCOPA, became a mediational means for Kris to perform the difficult cognitive task of producing the appropriate word order in real time speech.

As a result of CBI, the learners in Zhang's (2014) study not only were able to develop the ability to process Chinese word order without adhering to the constraints proposed by Processability Theory and the Teachability Hypothesis, the study also provided suggestive evidence that because of external forms of mediation (i.e., Cuisenaire rods) one learner was able to overcome an assessed deficit in her working memory. Of course, it is not possible to generalize from one case that all learners are likely to benefit from the use of external mediation, but it is a topic that is worth pursuing in the future, given that a good deal of L2 research has argued that working memory has an important influence in shaping learning outcomes (see Williams, 2012).

Pedagogical Implications

While we believe that the pedagogical implications of CBI as outlined and exemplified in this chapter are transparent, we nevertheless would like to highlight what we see as some differences between developmental education and other approaches to L2 instruction. For one thing, unlike other approaches, CBI considers development to take place not only at the level of concrete performance but also at the level of learner understanding of the concepts that underlie performance. In keeping with Vygotsky's notion of development, this knowledge provides learners with greater flexibility in generalizing performance across a wide array of contexts and thus enables them to use language in more creative ways (see Yáñez-Prieto, 2014). Furthermore, CBI, because it relies on conceptual knowledge, privileges theories of language such as cognitive linguistics and systemic functional linguistics, that foreground meaning rather than structure. Concomitantly, the approach places a good deal of responsibility on teachers to formulate explanations and ways of visualizing/materializing these explanations in pedagogically effective configurations. There is no single way of doing this. Much depends on the level and background of one's students. However, the nature of the linguistic concepts that are the object of instruction cannot be compromised. Often teachers feel that concepts such as pragmatic word order in Chinese, temporal aspect and mood in Spanish, phrasal verbs and verb + noun collocations in English, and so forth may be too complex for their students and therefore they adopt a piecemeal manner of presentation. As the research carried out by the Gal'perin team discovered (see Talyzina, 1981), however, eliminating any phase of CBI, including the formulation of complete concepts, has a deleterious effect on student development. Concepts must remain intact, but the way of visualizing/materializing them may vary depending on the nature of the students. For instance, Karpova (1977) used different colored plates to explain and illustrate word order variation to young children.

CBI clearly diverges from discovery learning, whose advocates argue that it is more effective for learners to explore the object of study in order to induce through either guided or unguided observation of a rule or a concept, in a way that parallels what

James P. Lantolf and Xian Zhang

occurs in the everyday world (see Egan, 2002; Wells, 1999). The difficulty with such an approach to education is that it is generally time-consuming and, more importantly, often results in incomplete or inaccurate learning outcomes (Karpov, 2014). CBI, on the other hand, relies on a conceptual foundation formulated through rigorous scientific research. Even if the findings of this research are tentative and open to revision, this knowledge is nevertheless superior both to observation in a limited array of empirical contexts and to commonsense reasoning. Having argued that CBI privileges scientific knowledge, it in no way devalues experiential learning. However, the experimentation that learners engage in is not about figuring out the nature of a concept; this information is provided by instruction. It is rather about learners having the opportunity to explore how to effectively use this knowledge for their own (communicative) purposes. This phase of education can be carried out under the guidance of a teacher or in cooperation with peers. Said in another way, theoretical/conceptual knowledge is of little value to learners unless they can link it to practical activity. Without that link, instruction would result in intellectualism, or what Vygotsky (1987) often labeled “verbalism.” On the other hand, performance that is not guided by high quality conceptual knowledge results in “mindless” behavior (Vygotsky, 1987).

Future Directions

Future work on CBI can be expected to focus on three general areas of concern: (1) extending CBI to languages beyond the current set; (2) continuing to broaden the domains of language that have been the topic of instructional interest to include, among other things, the ability to use and comprehend figurative language (i.e., metaphor and metonymy), to express motion events (i.e., manner and path of motion), and to express and comprehend emotion; and (3) preparing teachers to adapt and implement CBI procedures in their own educational environments. To our knowledge, CBI has been restricted to a relatively small set of languages, including Spanish, French, Chinese, Korean, and English. It is important to add to this number, because for one thing, different languages cover different linguistic concepts. It is therefore necessary to test the effectiveness of this approach to language education with as broad an array of concepts as possible. This is, of course, related to the second area—broaden the scope of the concepts addressed in CBI, including in the set of languages just mentioned. To date, grammar has been the primary focus of interest; however, in addition to Zhang’s (2014) study, some other work has been carried out on pragmatics (see Kim, 2013; van Compernelle, 2014), and while two studies have addressed literacy (Buescher, 2015; Ferreira, 2005), much more work needs to be dedicated to this topic. At least one instructional program has concerned itself with instruction on the semantics of motion events and the connection between speaking and gesturing (see Lantolf, Stam, Buescher, & Smotrova, 2014). Instruction on vocabulary and phonology has not been considered at all. Finally, L2 researchers have begun to turn their attention to the importance of communicating emotion through a new language—an area of concern that did not go unnoticed by Vygotsky and his colleagues (see Vygotsky, 1987). Emotional concepts, usually conceptualized through metaphor (e.g., he is so angry, steam is coming out of his ears), should be an especially provocative topic for CBI to address.

While it is all well and good to provide an exegesis on a particular approach to language education in the pages of a journal or an edited volume, it is quite another for someone to be able to adopt, and adapt, that approach solely as a result of reading about it. Thus, the third topic that we propose for future CBI research is to work with teachers and teacher educators

to help them understand and appreciate fully the psychological principles that underlie the approach and to experiment with CBI in their own instructional practices (see van Compernelle & Henery, 2015).

References

- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.
- Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189–208.
- Baddeley, A. D. (2010). Working memory. *Current Biology*, 20, 136–140.
- Bonilla, C. (2012). *Testing processability theory in L2 Spanish: Can readiness or markedness predict development?* (Unpublished doctoral dissertation). University of Pittsburgh.
- Buescher, K. (2015). *Developing narrative literacy in a second language through concept-based instruction and a division-of-labor pedagogy* (Unpublished doctoral dissertation). Pennsylvania State University.
- Cole, M. (2009). The perils of translation: A first step in reconsidering Vygotsky's theory of development in relation to formal education. *Mind, Culture, Activity: An International Journal*, 16, 291–295.
- Davydov, V. V. (2004). *Problems of developmental instruction: A theoretical and experimental psychology study*. Translated by P. Moxhay. Moscow: Akademiya.
- Di Pietro, R. (1987). *Strategic interaction*. Cambridge: Cambridge University Press.
- Egan, K. (2002). *Getting it wrong from the beginning: Our progressivist inheritance from Herbert Spencer, John Dewey, and Jean Piaget*. New Haven, CT: Yale University Press.
- Ferreira, M. (2005). *A concept-based approach to writing instruction: From the abstract concept to concrete performance* (Unpublished doctoral dissertation). Pennsylvania State University.
- Gal'perin, P. Ya. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology* (pp. 249–272). New York: Basic Books.
- Gal'perin, P. Ya. (1979). The role of orientation in thought. *Soviet Psychology*, 18, 19–45.
- Gal'perin, P. Ia. (1989a). Mental actions as a basis for the formation of thoughts and images. *Soviet Psychology*, 27(2), 45–64.
- Gal'perin, P. Ia. (1989b). Organization of mental activity and the effectiveness of learning. *Soviet Psychology*, 27(2), 65–82.
- Gal'perin, P. Ia. (1992). State-by-stage formation as a method of psychological investigation. *Journal of Russian and East European Psychology*, 30(4), 60–80.
- Gathercole, S. E., Tiffany, C., Briscoe, J., & Thorn, A. (2005). Developmental consequences of poor phonological short-term memory function in childhood: A longitudinal study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 46, 598–611.
- Gattegno, C. (1963). *Teaching foreign languages in schools: The silent way*. New York: Educational Solutions.
- Haenen, J. (1996). *Piotr Gal'perin: Psychologist in Vygotsky's footsteps*. New York: Nova Science.
- Haenen, J. (2001). Outlining the teaching-learning process: Piotr Gal'perin's contribution. *Learning and Instruction*, 11, 157–170.
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, 12, 9–15.
- Karpov, Y. V. (2014). *Vygotsky for educators*. New York: Cambridge University Press.
- Karpova, S. N. (1977). *The realization of language in children*. Paris: Mouton.
- Kim, J. (2013). *Developing conceptual understanding of sarcasm in a second language through concept-based instruction* (Unpublished doctoral dissertation). Pennsylvania State University.
- Kinard, J. T., Sr., & Kozulin, A. (2008). *Rigorous mathematical thinking: Conceptual formation in the mathematics classroom*. New York: Cambridge University Press.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, 24, 781–791.

James P. Lantolf and Xian Zhang

- Krashen, S. (2000). *The natural approach: Language acquisition in the classroom*. London: Pearson Education Limited.
- Lantolf, J. P. (2011). Integrating sociocultural theory and cognitive linguistics in the second language classroom. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning vol. II* (pp. 303–318). New York: Routledge.
- Lantolf, J. P., & Beckett, T. (2009). Research timeline for sociocultural theory and second language acquisition. *Language Teaching*, 45, 459–475.
- Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. New York: Routledge.
- Lantolf, J. P., & Poehner, M. E. (Eds.). (2008). *Sociocultural theory and the teaching of second languages*. London: Equinox.
- Lantolf, J. P., Stam, G., Buescher, K., & Smotrova, T. (2014, October). *Teaching thinking for speaking*. Paper presented at the annual meeting of the Sociocultural Theory and Second Language Research Working Group, Coral Gables, FL.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.
- Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory and second language learning. In B. VanPatten & J. Williams (Eds.), *Theories in second language learning: An introduction* (pp. 201–224). Mahwah, NJ: Lawrence Erlbaum.
- Lee, H. (2012). *Concept-based approach to second language teaching and learning: Cognitive linguistics-inspired instruction of English phrasal verbs* (Unpublished PhD dissertation). Pennsylvania State University.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- McNeill, D. (2005). *Gesture & thought*. Chicago: University of Chicago Press.
- Miller, R. (2011). *Vygotsky in perspective*. New York: Cambridge University Press.
- Morgan, P. S. (1997). Figuring out figure out: Metaphor and the semantics of the English verb-particle construction. *Cognitive Linguistics*, 8, 327–357.
- Negueruela, E. A. (2003). *A sociocultural approach to the teaching-learning of second languages: Systemic-theoretical instruction and L2 development* (Unpublished doctoral dissertation). Pennsylvania State University.
- Ortega, L. (2007). Second language learning explained? SLA across nine contemporary theories. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 225–250). Mahwah, NJ: Lawrence Erlbaum.
- Ortega, L. (2015). Second language learning explained? SLA across 10 contemporary theories. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 245–272). New York: Routledge.
- Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, 10, 52–79.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.
- Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23, 193–229.
- Swain, M. (2006). Linguaging, agency and collaboration in advanced language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). Washington, DC: Georgetown University Press.
- Talyzina, N. (1981). *The psychology of learning: Theories of learning and programed instruction*. Moscow: Progress Press.
- Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. New York: Routledge.

- Valsiner, J. (1997). Magical phrases: Human development, and psychological ontology. In C. Lightfoot & B. D. Cox (Eds.), *Sociogenetic perspectives on internalization* (pp. 237–256). Mahwah, NJ: Lawrence Erlbaum.
- van Compernelle, R.A. (2012). *Developing sociopragmatic capacity in a second language through concept-based instruction* (Unpublished doctoral dissertation). Pennsylvania State University, University Park.
- van Compernelle, R.A. (2014). *Sociocultural theory and L2 instructional pragmatics*. Bristol: Multilingual Matters.
- van Compernelle, R.A., & Henery, A. (2015). Learning to do concept-based pragmatics instruction: Teacher development and L2 pedagogical content knowledge. *Language Teaching Research*, 19, 351–372.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky volume 1: Problems of general psychology: Including the volume thinking and speech*. Edited by R. W. Rieber & A. S. Carton. New York: Plenum.
- Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.
- Williams, J. N. (2012). Working memory and SLA. In S. M. Gass & M. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 427–441). New York: Routledge.
- Yáñez-Prieto, M. C. (2014). *Sense and subjectivity: Teaching literature from a sociocultural Perspective*. *Language and Sociocultural Theory*, 1(2), 179–204.
- Zhang, X. (2014). *Teachability and topicalization in L2 Chinese* (Unpublished doctoral dissertation). Pennsylvania State University.
- Zhang, X., & Lantolf, J. P. (2015). Natural or artificial: Is the route of second language development teachable? *Language Learning*, 65(1), 152–180.

Processing Instruction

Bill VanPatten

Background

This chapter begins with a clarification about Processing Instruction. Processing Instruction (PI) is sometimes taken to be an approach (or method), but it is not. *Approach* is an overall theoretical framework for a method that includes a theory of language and a theory of language acquisition (e.g., Richards & Rogers, 2001). For this reason, there are communicative approaches, functional approaches, and proficiency-based approaches, to name some of the broader categories. From an approach, both design and procedure are derived to develop a particular method (e.g., within communicative approaches there are different kinds of immersion, there is the Natural Approach, there is TPRS). As will be seen, PI is actually a type of focus on form or better yet, a pedagogical intervention. As such it is not a method with an underlying approach but instead an intervention that can be used by any communicative approach that seeks a supplemental or periodic focus on the formal features of language. However, like many other pedagogical interventions, to understand PI requires the exploration of a number of topics that inform the nature and intent of PI: what is acquired (i.e., the nature of language), what acquisition is, what input processing is, and what these imply for a pedagogical intervention.

Language

Scholars working with PI understand that language is a complex, abstract, and implicit mental representation that cannot be captured with the simple rules that are typical of textbooks and much of instructed SLA research. I have argued elsewhere for a generative perspective on language (e.g., VanPatten, 1996, 2013; VanPatten & Rothman, 2014, 2015). Under this perspective, there are no rules or paradigms in a conventional sense and what we call a “sentence” is a surface manifestation of an underlying complex interaction between features, syntactic operations, the lexicon, and other components of language. The example I have offered before is the prototypical “rule” for the formation of *yes/no* questions in languages like English and Spanish. In English, the traditional or conventional rule is something like “insert *do* and invert with the subject” and in Spanish it is something like “invert the subject

and verb” (Spanish does not have an auxiliary *do* as does English). My argument is that such statements used to teach and practice language are psychologically unreal. The difference between statements and *yes/no* questions in the two languages in question is far more abstract than what meets the eye and involves (minimally) (1) differences in the nature of lexical verbs (Spanish lexical verbs contain the abstract feature T—Tense—whereas English lexical verbs do not, but auxiliaries do), and (2) abstract features in the CP (Complementizer Phrase) of questions that need to be satisfied in sentence structure. Spanish satisfies these abstract features one way (moving lexical verbs out of the VP and up into CP via the TP [Tense Phrase]) while English satisfies them another way (inserting a *do* into CP). Conventional or traditional rules of the type “insert *do* and invert with the subject,” then, are shorthand ways to describe something too abstract and complex to describe easily.¹ Such rules cannot be the starting point or object of acquisition.

So the first point toward understanding the intent of PI is that it is not like other interventions that claim or assume that acquisition is the internalization of rules. Instead, what learners acquire is an abstract, implicit, and interactive mental representation. Unfortunately, this chapter precludes a detailed discussion of language and so the reader may consult the publications previously cited (esp. VanPatten & Rothman, 2014) as well as Gregg (1989), Lardiere (2012), Slabakova (2012), and L. White (2015), among others.

One final point needs to be clear; that language and communication are not synonymous. Language is representation, but communication is a process that makes use of language (in humans). Communication is bound up in skill, social interaction, and other nonlanguage (but interconnected) aspects of human behavior. Thus, it is critical to keep in mind that the aim of PI is not to affect communication or skill but mental representation. (For the distinction between language and communication/skill, see VanPatten, 2013, 2016a, 2016b).

The Basic Nature of Acquisition

Under PI, acquisition consists of three necessary ingredients (putting aside social issues and context for discussion here): (1) input; (2) Universal Grammar (UG) and internal mental architecture; and (3) processing mechanisms that mediate between input and UG/internal architecture.

Input for acquisition is language that learners hear or see in a communicative context that they process for meaning (propositional content and intent). This definition excludes what many instructors believe to be key ingredients for acquisition such as explicit information and practice aimed at a particular feature. That is, explicit information about, say, how past tense works is not linguistic input for past tense. Only samples of past tense used in communicative contexts can serve as input for the acquisition of past tense (e.g., hearing someone say “I failed my test” and “I wrecked my car” as they talk about all the things that went wrong this week). Whether or not explicit information and practice play any effect on acquisition is irrelevant here. The point is that such things are not input for acquisition; that is, they are not the data that the internal mechanisms use to create and recreate language.

Internally, learners possess both UG and general learning mechanisms that interact with incoming data. Universal Grammar provides the language specific constraints so that the mental representation conforms to a human language. Under Minimalist Theory, these constraints include (1) a preset inventory of possible features (e.g., Tense, Case, Number) from which languages may select; (2) primitives or lexical and functional categories such as N(oun), V(erb), P(reposition), D(eterminer), and so on; (3) basic operations (e.g., Merge/Move, Agree/Check); and (4) a set of universal constraints (e.g., the Extended Projection

Bill VanPatten

Principle, the Overt Pronoun Constraint, the Locality Principle). General learning mechanisms assist the language learner in figuring out meaning (e.g., What does this word mean? What is this person saying to me?).

Universal Grammar and the internal mechanisms do not (typically) operate directly on input data. Instead, there is some kind of processing of input data that converts it into something usable by UG and the internal mechanisms. This processing may also filter input data, resulting in a subset of the data called *intake*. Input processing, then, is a kind of buffer between the “possible data out there” and the “mechanisms waiting for data in the head.” Thus, before UG and the internal mechanisms can operate on any input, that input first passes through a mechanism that processes and tags it in particular ways.

I have sometimes used the following analogy to illustrate the three essential components of acquisition. In contemporary supermarkets, when one checks out at the register, three things are minimally necessary for the creation of a total cost: barcodes on products, a computer with its internal workings, and an infrared scanner. The barcode (not the labels, not the pictures on the bag or can, not anything else) is the input required to register a price. The computer (UG/internal mechanisms) assembles all of the prices to create a total cost (mental representation). The infrared scanner (input processor) is the buffer between the barcodes and the computer. That is, the internal workings of the computer do not read the barcodes directly; the infrared scanner does this and converts the barcode into something usable by the computer. For supermarket checkouts, then, all three ingredients are necessary: barcodes, infrared scanners, and internal computer. Likewise, for the creation of a mental representation of language, the three ingredients of input, input processors, and UG/internal mechanisms are all necessary. None of them can be left out of the process of acquisition and none can be substituted by something else (e.g., explicit information and practice as noted earlier).²

Input Processing

Critical to understanding PI is an understanding of input processing (“how the scanner works”). Two points are necessary in this regard. The first focuses on what gets processed. The second refers to basic principles that constrain or guide input processing. We will take each in turn.

If we take as our premise that there are no rules “out there” to internalize, then what do learners process in the input? That is, if learners’ processors are *not* “scanning the input” for the formation of *yes/no* questions or rules for passive formation or rules for how to create the simple past tense, then what do the input processors focus on? Here I quote directly from VanPatten and Rothman (2014) as we summarize our position on this matter:

Learners do not acquire rules from the input. Instead, learners process surface morpho-phonological units (e.g. lexical form, morphological form) and internalize these units along with underlying features or specifications. These units interact with information provided by UG and the language making mechanisms of the human language faculty such that anything that resembles rules (from an outside perspective) evolves over time.

p. 25

The main point to be taken from this quote is that what learners get from the input are lexical and lexical-like pieces of language (e.g., words and morphological properties of words). As learners’ processors engage input, they do not process *yes/no* questions, for example,

as having a rule but instead process *do/does/did* as lexical units along with their underlying features and their serial position in an utterance. So, for example, they must eventually process *does* and store it with the following features: [-N], [+V], [+T], [+present], [-past], [-1st person], [-2nd person], [-plural], among, possibly, others. Likewise in Spanish, learners' processors do not process *yes/no* questions as having a rule but instead must eventually process lexical verbs such as *tomas* ("you drink") as having the following features: [meaning of *tom-*], [+T], [-N], [+V], [+present], [-past], [-1st person], [+2nd person], [-plural], and so on. It is this information that allows the verbs in their respective languages to enter into syntactic operations to yield the surface *yes/no* questions that we see.³

Against this nondetailed backdrop, we also need an understanding of how lexical and lexical-like information gets processed as intake data, including how learners compute basic relationships among elements of a sentence. In various publications, I have sketched out a general model of input processing (e.g., VanPatten, 1996, 2004, 2015a) that includes three basic principles of input processing (and their corollaries) to capture basic strategies that guide learners' processors: the Primacy of Meaning Principle, the Lexical Preference Principle, and the First-Noun Principle.

The Primacy of Meaning Principle states that learners' comprehension of input is driven by a focus on meaning and not a focus on form. That is, learners do not consciously or unconsciously approach comprehension with the intent of seeking out formal properties of language. Instead, their intent is to figure out the meaning of what they hear or see. Because L2 learners are not like L1 learners who have to discover that lexical units and lexical phrases exist, they come to the task knowing that there is a word for a four-legged feline or that there is a way to greet someone in the hallway, for example. Thus, a natural consequence of the Primacy of Meaning Principle is that learners seek out content lexical items and lexical phrases from the beginning as building blocks to meaning. This search for and isolation of lexical items and phrases may be aided by prosody, context, and interaction; however, the point here is that internally, the learner is working at building up a lexicon.

The initial focus on lexical items and phrases as key to meaning suggests a second major principle: Lexical Preference. The Lexical Preference Principle says that, assuming learners must process morphological information at some point (such as past tense markings), they initially process lexical items as cues to the underlying meaning of morphological features. Concretely, in the case of something like the past tense, learners will first seek out and process words and phrases that are tense indicators (e.g., *right now*, *everyday*, *last week*, *next year*) to comprehend the temporal reference of an utterance or discourse. The idea underlying this principle is that most grammatical inflections are redundant in nature; that is, markers of such things as plurality, person and number, tense, and so on, are almost always recoverable in the input from lexical information (e.g., *two*, *many* = plural; *John*, *he* = third person singular, *yesterday* = past). Because the underlying features of morphological inflections are generally engaged in some kind of Agree relationship in the syntax, the Lexical Preference Principle predicts that learners don't separately process and tag morphological inflections until they have enough robustly represented lexical items in their inventory with which to match them. That is, *watched*, *talked*, and *typed*, as three examples, are not processed as having the underlying features [-present] [+past] until there are enough adverbials stored in the lexicon that carry the same feature (e.g., *yesterday*, *last week*, *last night*, *2 weeks ago*). In developmental time, this means that adverbials enter the lexicon before inflections related to temporal features are coded on verbal lexical entries.

Another basic principle of input processing is the First-Noun Principle, which states that learners tend to process the first noun (or pronoun) they encounter as the subject/agent of

a sentence. Although this principle works well for basic SVO type sentences, it impedes the correct processing of passives in a language like English, for example, and non-SVO sentences in languages like Spanish, case marking in languages like German and Russian, among other formal surface features of language. When learners incorrectly process reversible passives such as *The man was followed by the woman* as “The man followed the woman,” critical data is missing for the development of mental representation. When learners incorrectly process sentences such as *Lo ve la profesora* “The professor sees him” as “He sees the professor,” they are delivering faulty data to the internal mechanisms responsible for the development of mental representation. And when learners rely on the First-Noun Principle to process *Die Frau hört den Mann* “The woman hears the man,” they do not make use of case marking to determine who does what to whom, and data may be filtered out of the input as it is delivered to the internal mechanisms.

These principles all have corollaries that explicate and expand on how they function and interact, and the reader is referred to other publications for more detailed information (e.g., Farley, 2005; VanPatten, 1996, 2004). Before continuing, it is important to stop and examine the term *process*. In the model of input processing outlined here, the term “process” refers to how learners connect form with meaning during the act of comprehension. This definition is critical because input processing cannot be equated with something like “noticing,” which underlies other pedagogical interventions (e.g., text enhancement, recasts). Noticing is a term with some elasticity and, like many constructs, may be used differently by different scholars. As originally defined by Schmidt (1990, and elsewhere), noticing is essentially some conscious registration of something new in the linguistic input. Conscious registration means that learners become aware of something they had not been aware of before. For example, a learner might hear “talked” and realize that it is different from either “talk” or “talking.” That learner has noticed it. Schmidt is clear that noticing does not entail any kind of awareness of what is noticed; that is, noticing does not mean linking form with meaning. Thus, noticing and processing are not synonymous (see VanPatten, 2014, 2015b, as well as Truscott, 1998, for further discussion). This distinction is critical for the reader to understand the intent of PI. Unlike other pedagogical interventions, it is not the intent of PI to “induce noticing” or to make something in the input “salient.” We will return to this point later.

To conclude this brief overview of input processing, the most significant (and yet deceptively simple) point is this: that learners bring to the task of acquisition a set of principles that guides how they link form with meaning during comprehension. These principles constrain the nature of the data that is delivered to the internal mechanisms. In short, input processing is a “bottleneck” where rich input data is culled to deliver intake data to UG and the internal mechanisms responsible for creating a linguistic system. Because this processing reduces the data and/or alters it (i.e., processes it incorrectly), the acquisition of mental representation is less than optimal. The question then becomes whether it is possible to assist learners during input processing such that their processing is altered. If so, altering input processing subsequently alters the quantity and quality of the intake data, thus enriching the acquisition of mental representation. This idea was articulated in the foundational study of PI, VanPatten and Cadierno (1993): “Input processing is concerned with [. . .] the conversion of input to intake” (p. 226) and “Theoretically, altering input processing should have a significant impact on changing internalized knowledge” (pp. 227–228).

To underscore something critical here, the reader’s attention is drawn to this distinction: the focus of PI is not helping learners uncover rules or paradigms in the input; instead, the focus of PI is helping learners correctly process morpho-phonological units to enrich the intake provided to the internal mechanisms responsible for the creation of a mental

representation of language (see VanPatten & Rothman, 2014, cited earlier, as well as VanPatten & Rothman, 2015).

Pedagogical Implications

The implications of the previous sketch—and in particular the brief outline of input processing—should be clear. Typical practice scenarios in much of language teaching involve the tail wagging the dog. That is, teachers explain rules or forms (or learners read about them), and then learners practice them in controlled production activities (either meaningfully or nonmeaningfully). Yet, the creation of mental representation is input dependent. Thus, having learners “practice grammar” attempts to obviate the role of input in acquisition. Such attempts just don’t fare well (for some early and classic research on this, see Lightbown, 1983).

At the same time, pedagogical interventions such as text enhancement, input flood, dictogloss, and many others do recognize the role of input in language acquisition. However, they are either ill-informed or underinformed regarding the nature of language, the nature of acquisition, and the nature of input processing. The central idea put forth in this chapter is that for a pedagogical intervention to be useful in the creation of a mental representation of language, it must:⁴

- Clearly delineate what it believes the learner is acquiring (i.e., what will wind up in the mind/brain);
- Clearly lay out the minimal ingredients and mechanisms involved in language acquisition;
- Have some firm description of the nature of how learners’ processors deal with raw input data under uninstructed conditions.

From this perspective, one implication is that interventions that help with the creation of mental representation ought to be processing-oriented pedagogical interventions—or POPIs for short.⁵ Such interventions claim that the first step (but not the only) in developing a representation of language involves the processing of input data from the environment. Once again, processing means that some mechanism isolates a morpho-phonological unit in the input stream and attaches both a meaning and a function to it; in short, form and meaning are linked both at the local level (e.g., the word/form) and the sentence level. POPIs are not predicated on noticing (see earlier discussion of the distinction between processing and noticing). But what might a POPI look like? One such POPI is PI, the focus of this chapter.⁶

In PI, activities manipulate input so that the learners are forced to abandon the strategies embodied in the various principles of the model of input processing sketched previously. This manipulated input is referred to as *structured input*. *Referential* structured input activities within PI usually begin the intervention and are structured to have right or wrong answers. We can illustrate using the First-Noun Principle and its intersection with the processing of clitic object pronouns in Spanish—but the same PI activities can be used to inform work on case marking in languages like German and Russian (see VanPatten, Borst, Collopy, Qualin, & Price, 2013). As noted earlier, learners generally have success processing SVO sequences with the First-Noun Principle but not OVS and other sequences. Thus, they can correctly process the essential morpho-phonological units of *María ve al chico* “Mary sees the boy” but may skip the clitic object pronoun *lo* in *María lo ve* “Mary sees him” or jettison it during processing because it cannot be connected to meaning (it appears before the

verb but *María* has already been tagged as the subject). Alternatively, they may misinterpret *Lo ve María* as “He sees Mary” incorrectly tagging the clitic object pronoun as a subject pronoun. In a series of referential structured input activities in a PI sequence, learners hear a mixture of SVO/SOV, OVS, OV sentences in which both the subject and object are capable of performing the action (e.g., a boy looking for a girl or a girl looking for a boy). Learners are asked to select between two pictures in order to indicate they have correctly processed and comprehended the sentence—or they may be asked in some other way to indicate who did what to whom. Such activities are designed to force the learners’ internal processors to abandon a strict reliance on the First-Noun Principle. For Spanish clitic object pronouns, then, this means correctly processing something like *lo* in *Lo ve María* as an accusative pronoun meaning “him” and interpreting the sentence as “Mary sees him” and not incorrectly as “He sees Mary.” Note that we are not making any claims about what is learned or that rules are being internalized. We are only stating here that learners are correctly tagging *lo* with its meaning and function in the utterance and over time will internalize its underlying features (e.g., [+accusative], [–fem], [–plural], [+3rd person], and so on). The continued correct processing of *lo* as a lexical item, then, is what strengthens it within the mental lexicon and allows it to participate subsequently in sentence structure.

Traditionally, referential activities are followed by *affective* activities. Unlike referential activities, affective activities do not have right or wrong answers (at least, not answers that are known by the learner) and focus on opinions, conclusions, personal experience, and so on. For example, in the string of PI activities on the First-Noun Principle and clitic object pronouns in Spanish sketched out earlier, an affective activity might be one in which learners select a female family member and must indicate how that person feels about her by checking a box as follows (translations provided for the reader, not for the learner):

- ☐ La respeto (“I respect her”);
- ☐ La detesto (“I hate her”);
- ☐ La comprendo bien (“I understand her well/I get her”)

and five other items.

In such an activity, learners find out about someone’s relationship with a particular family member, and then this activity may be repeated with a male relative with the exact same items (e.g., *Lo respeto*, *Lo detesto*, *Lo comprendo bien*) with the idea that learners will determine with whom the person has a better relationship. The purpose of affective activities is to provide a communicative context in which the continued correct processing instantiated by referential activities can occur.

We can also illustrate PI with something like morphological inflections using a simple past tense marker and its intersection with the Lexical Preference Principle. Based on the idea that learners use lexical items to make temporal reference assignments during comprehension, verbs may not be tagged for temporal features for some time during acquisition. PI as an intervention would consist of activities in which learners could not rely on lexical items such as adverbials of time to assign general temporal reference to sentences while listening or reading. A basic referential activity might involve hearing adverbial-less sentences with a mixture of temporal references (e.g., *John attends class*, *Mary talked on the phone*) and then selecting words that match the sentence (e.g., *yesterday* vs. *everyday* vs. *tomorrow*). The underlying motivation in this intervention is to push learners away from relying on adverbials to grasp temporal reference and instead rely on information in verbal cues as indicators of temporal reference. As learners process more and more verbs, they internalize the features associated with them so that such verbs can participate in sentence structure (e.g., *tomé* (“I

drank”): [meaning of “drink”], [+T], [−N], [+V], [−present], [+past], [+perfective], [+1st person], [−plural], and so on). As before, it is important to underscore that the intent of PI is not to teach verb forms or paradigms but instead to get learners to correctly process verbs with their encoded meanings (which include features) in order to link form and meaning. It is the internal processors (UG and the learning mechanisms) that take care of how mental representation develops.

Particular guidelines for the development of structured input activities have been provided in Lee and VanPatten (1995, 2003) and VanPatten (1993) and elsewhere. In addition, published scholarship is replete with extensive descriptions and examples of PI activities and interventions. For the reader’s convenience, the following publications are suggested: Farley (2005), VanPatten (1996), VanPatten et al. (2013), and Wong (2005). Farley’s book is a particularly good resource. It is imperative that if PI is to function, the treatment must adhere to particular guidelines, a point we will return to in the next section.

Key Concepts

Intake: This term refers to the subset of input data that the learner actually processes (see Processing) during the act of comprehension.

Processing: This term refers to learners linking form and meaning during real-time comprehension. It is not synonymous with the concept of noticing.

Mental Representation: This construct relates to the linguistic system in the mind/brain. It is an abstract, implicit, and complex system that does not consist of rules as classically or traditionally conceived.

Principles of Input Processing: These principles are a broad set of four major principles with corollaries that describe how the linking of form and meaning is constrained or filtered by L2 learners.

Referential Activities: These activities are found in PI and have right or wrong answers. They are the way in which PI typically begins as an intervention and are meant to push learners away from less-than-optimal strategies for making form–meaning links.

Affective Activities: These are PI activities that typically follow referential activities. They do not have right or wrong answers and provide additional structured input to allow learners to continue making form–meaning links they began under referential activities.

Empirical Evidence

Within the scholarship on pedagogical interventions and focus on form techniques available (see, for example, Doughty & Williams, 1998), PI is one of the most widely researched. At the same time, it is the intervention with the most robust results in terms of effects, while also studied in the most contexts, with the most languages, with a variety of intersections of processing problems and target forms. Since the foundational studies of VanPatten and Cadierno (1993) that focused on the First-Noun Principle, and Cadierno (1995) that focused on the Lexical Preference Principle, research on PI can be found in over 40 studies. Following is a nonexhaustive list of published research and the general findings.

- *The role of explicit information and explicit feedback.* Fernández (2008), Henry, Culman, and VanPatten (2009), Sanz and Morgan-Short (2004), VanPatten et al. (2013),

VanPatten and Oikonen (1996), and J. White and DeMil (2014). These studies have found that providing learners with explicit information prior to the intervention is not necessary; that is, whether or not learners receive explicit information in PI is not as important as the structured input activities that seem to push processing in appropriate directions.

- *The use of different assessment measures/transfer of training.* Henry (2015), Sanz and Morgan-Short (2004), VanPatten and Sanz (1995), VanPatten and Uludag (2011), and J. White and DeMil (2014). The foundational studies on PI used sentence-level assessment tasks focused on correct sentence interpretation. Research since then has manipulated assessment tasks to include story narration and text reconstruction. All such manipulations have yielded effects for PI. At the same time, every study consistently shows effects on the gold-standard of PI research: sentence-level processing tasks.
- *The role of aptitude/individual differences.* Lee and Benati (2013) and VanPatten et al. (2013). Emerging from these studies is that the results of PI do not correlate with or depend on individual differences such as aptitude (as traditionally measured by such tests as the Modern Language Aptitude Test) or working memory.
- *PI and discourse level effects.* Benati and Lee (2012). Research falling into this group has demonstrated that although PI as in intervention is confined largely to sentence-level activities and that most interpretation assessments are sentence level, the effects of PI can be seen on discourse-level tasks (e.g., interpretation tasks that do not involve isolated and decontextualized sentences).
- *Secondary effects.* Benati and Lee (2008) and J. White and DeMil (2013). The studies in this line of research suggest that the effects of PI may transfer to the processing of morpho-phonological units that are not the object of the PI itself. For example, the correct processing of clitic object pronouns shows some effects on the processing of dative clitics even when the latter are not the target of intervention.
- *Long-term effects.* VanPatten and Fernández (2004). In this singular study, the effects of PI were still evident almost nine months after treatment, albeit with some decline. That is, there was decline from an immediate posttest to a delayed posttest, but the scores on the delayed posttest were significantly greater than those on the pretest.
- *Comparisons with other interventions.* Benati (2001, 2005), Cadierno (1995), Comer and deBenedette (2010), Uludag and VanPatten (2012), VanPatten and Cadierno (1993), VanPatten, Farmer, and Clardy (2009), VanPatten, Inclezan, Salazar, and Farley (2009), and J. White, DeMil, and Rice (2015), and many others. This group of studies has formed the bulk of early PI research and even some of the PI research into the 2010s. The effects of PI have been compared to what can be called traditional teaching of grammar (e.g., explanation plus practice), dictogloss, meaning-based output instruction, and others. In each and every case, PI outperforms others on tests of interpretation and processing and is equal or better on other kinds of tasks.

To be sure, there have been some detractors who have attempted to discount the research on PI (e.g., DeKeyser & Prieto Botana, 2015; DeKeyser, Salaberry, Robinson, & Harrington, 2002; DeKeyser & Sokalski, 1996), but as argued in various responses and other publications, the objections raised are due to fundamental misunderstandings about the nature of PI (see, Sanz & VanPatten, 1998; VanPatten, 2002, 2015b; Wong, 2004; among others). For example, DeKeyser and Sokalski reduce PI to mere “comprehension.” In their purported replication research their misunderstanding is clear as their treatment bears no resemblance to anything like a POPI as described here (and as described in research as far back as

VanPatten & Cadierno, 1993). In addition, the objections raised in the citations noted earlier often reveal a belief in the existence of conventional rules and that acquisition involves internalizing rules. Thus, the research on PI is viewed from an inappropriate lens; that is, apples are being compared to oranges.

Others have claimed to find results that do not mirror those of PI research (e.g., Allen, 2000; Qin, 2008). However, in various responses that involve replication research, the problems in research design and treatment of those studies are remedied and the results dovetail with the rest of PI research (e.g., Uludag & VanPatten, 2012; VanPatten & Wong, 2004). In short, claims for different results from standard PI research are traceable to problems in how those researchers create treatments and in the assessments they use. Here we come back to the guidelines for the creation of appropriate PI activities mentioned in the previous section. In studies such as Allen (2000), DeKeyser and Solkaski (1996), and Qin (2008), the basic guidelines for the creation of appropriate structured input activities were either ignored or misunderstood, resulting in treatments that naturally led to results different from those of standard PI studies.

Another problem in some of the critiques or discussions of PI are inappropriate outcome measures. For example, in Marsden and Chen (2011), nonprocessing and noninterpretative measures of “rule learning” were used, which resulted in their particular conclusions about the role of explicit learning in PI. Again, the standard of measurement in PI research should not be related to knowledge outcome but to the ability to process correctly and demonstrate form–meaning links during online comprehension.

Teaching Tips

- *Keep in mind the difference between representation and skill.* PI is not an approach or method for teaching communicative skill. It is an intervention for assisting in the development of mental representation of language.
- *Have clear expectations.* No pedagogical intervention that is a focus on form causes instant acquisition. It is important to keep in mind that acquisition of a new linguistic system is slow and piecemeal. An intervention like PI is not a magic bullet; it is an aid.
- *Guidelines.* There are guidelines for the development of PI activities that must be followed to ensure that one is developing an appropriate PI treatment. Perhaps the most important guideline is that the intervention must keep the processing strategies in mind; that is, the intervention must be constructed such that processing is actually altered because the input is structured in such a way to push processing into a different direction. Other guidelines include focusing on one thing at a time, moving from sentences to discourse (i.e., not beginning with discourse), and working with both oral and written input, among others.

Current Issues and Future Directions

As noted earlier, PI has enjoyed an extensive research agenda within the field of instructed SLA. Given the myriad of variables and factors within and related to PI that have been researched since 1993, it is difficult to imagine where else the research could head. Nonetheless, a few of the areas listed earlier offer some ideas.

Although research on PI and individual differences is reported, it is fair to say that this research is incipient and is just scratching the surface. As VanPatten et al. (2013) argue, there is good reason based on their research to conclude that aptitude as traditionally measured is not and should not be a major variable in the outcomes of PI. In their rather large study with four languages (Spanish, French, German, and Russian) with four different structures affected by the First-Noun Principle, aptitude scores simply did not correlate with or impact the outcomes in any of their experiments. Of particular importance is that one of the major assessments was trials-to-criterion or “how long it took a learner to begin to process correctly”—an item-by-item measure during the treatment phase. In short, aptitude did not emerge as a variable affecting processing in their study. At the same time, when one looks closely at the data, there is variation among participants as in most studies of instructed SLA. In VanPatten et al., four experiments were reported but we will examine data from just one here. In the Spanish experiment, 9 participants out of 42 did not reach criterion at the end of the treatment. What this means is that 9 participants when tracked item by item (this was a computer-based study) never evidenced the criteria for correct processing (i.e., three items plus a distractor plus 60% correct overall thereafter). However, 33 participants did. A closer scrutiny of the means and standard deviations show very wide deviations. In one of the Spanish subgroups, for example, the mean was 16.63 but the standard deviation was 17.17. These numbers suggest tremendous variation among the learners in terms of how the treatment affected them during the study. So while in this study the focus was on aptitude and the results of all four experiments reveal no significant role for aptitude as a variable, we are left with considerable individual difference in treatment effects that requires explanation. As VanPatten et al. argued, one should not expect a role for aptitude in PI because aptitude is about rule learning but PI is about processing morpho-phonological units in the input. This research suggests, then, that something like input processing and any effects of instruction on it, are sensitive to individual differences that we do not yet understand. The field is ripe, then, for looking into new formulations of aptitude unrelated to rule learning.

Another area of potential interest that has emerged from the research on PI that has not emerged from other areas (e.g., text enhancement, recasts) is that although different kinds of morpho-phonological units in the input ultimately respond favorably to a PI treatment, there seem to be differences in rates and differences in how explicit information interacts with the processing of a form. For example, in VanPatten et al. (2013), the processing problem was held constant across the four studies: the First-Noun Principle. However, four different structures were used across four languages: Spanish (word order and clitic object pronouns); German (case marking on articles); Russian (case marking on nouns); and French (causative construction with *faire*). Across the four studies, there seems to be differences in when “processing kicks in,” with the correct processing of Spanish clitic object pronoun structures coming in sooner than the causative in French (i.e., $M = 16.63$ for Spanish clitics and $M = 29.68$ for causatives; the reader is reminded that in this study the researchers were looking for how soon learners began to process correctly, so lower scores are better as they represent the mean item number at which learners began to process correctly). The question here, then, is why some structures interact with a particular processing problem differentially. In the case just presented, why do learners of Spanish begin to correctly process clitic object pronouns sooner than when learners of French begin to correctly process causative *faire*?

At the same time, such research suggests that explicit information plays a different role depending on the structure (when the processing problem is held constant). In the VanPatten et al. study, explicit information was found to be beneficial in the processing of German case and the French causative with *faire*, but not with Russian case or Spanish clitic object

pronouns. Why would explicit information play differential roles? Such a question leads to perhaps the most fundamental issue regarding explicit information during acquisition: Under what conditions can learners make use of explicit information during the processing of input strings in an L2?

Conclusion

In this chapter, the following main points about PI have been made:

1. PI is not an approach or method, but a pedagogical intervention to assist in the development of mental representation, not skill.
2. PI does not focus on rule or form learning but the correct processing of morpho-phonological units in the input in order to link form and meaning during real time comprehension.
3. PI is informed by a particular model of input processing and unlike other input-oriented pedagogical interventions is not predicated on the concept of “noticing” but instead on the concept of altering processing strategies.
4. There is a rich and robust research agenda on PI in which its impact on acquisition has been examined from a variety of perspectives; the result is over 40 studies on PI to date.
5. The research is unequivocal on the consistent positive impact of PI as measured by tests of processing and interpretation.
6. New directions in PI research include exploring novel ways of examining differences in individual performance (e.g., moving away from traditional notions of aptitude) as well as how target forms interact differentially with particular processing problems during treatment.

As this conclusion is written, it is the summer of 2015. The foundational study for PI appeared in VanPatten and Cadierno (1993). Thus, we are closing in on almost 25 years of research on this one pedagogical intervention. It is not clear that research on PI will end anytime in the near future. That PI has had an enduring research agenda speaks to its solid connection to a theory of language, a basic theory of acquisition, and a clearly delineated set of expectations about what PI should affect in acquisition and how it should be affected. That PI is a beneficial tool for a communicative curriculum is not in doubt and what remains to be seen is an accounting for individual differences in performance—and this is interesting because it clearly links PI to the broader field of second language acquisition. In short, PI may be a way to test various variables and issues in SLA more generally.

Notes

1. Although I take a generative perspective on the nature of language, PI is compatible with other approaches (e.g., emergentism/usage based theories, complexity theory) because these approaches also do not subscribe to rules in the conventional sense.
2. Positing these three ingredients as necessary does not mean that acquisition is guaranteed. Non-nativeness in various domains of language is the norm in SLA. At the same time, positing these three ingredients as necessary does not obviate the possible role of L1 influence.
3. From this scenario, it should be clear that inflections (e.g., nominal or verbal) are not acquired separately from lexical items but along with them. Learners internalize whole words, not parts of words. Inflections are later tagged with meaning as part of the larger word they occur with and what are typically called “productive inflections” are derived from the lexicon. They are not initially learned and stored as some separate component of the grammar.

Bill VanPatten

4. See also VanPatten (2015c).
5. The pronunciation of this acronym is POE-peeh. In addition, parts of this section of the chapter have been adapted from VanPatten (2015c).
6. Currently, there are no other POPIs in existence. That is, there are no POPIs derived from an understanding of the principles underlying input processing. For discussion related to this, see VanPatten (2009, 2015c).

References

- Allen, L. Q. (2000). Form–meaning connections and the French causative: An experiment in processing instruction. *Studies in Second Language Acquisition*, 22, 69–84.
- Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research*, 5, 95–127.
- Benati, A. (2005). The effects of processing instruction, traditional instruction, and meaning-output instruction on the acquisition of the English past simple tense. *Language Teaching Research*, 9, 87–113.
- Benati, A., & Lee, J. F. (Eds.). (2008). *Grammar acquisition and processing instruction: Secondary and cumulative effects*. Bristol: Multilingual Matters.
- Benati, A., & Lee, J. F. (Eds.). (2012). *Processing instruction and discourse*. London: Continuum.
- Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *Modern Language Journal*, 79, 179–193.
- Comer, W. J., & deBenedette, L. (2010). Processing instruction and Russian: Issues, materials, and preliminary experimental results. *Slavic and East European Journal*, 54, 118–146.
- DeKeyser, R. M., & Prieto Botana, G. (2015). The effectiveness of processing instruction in L2 grammar acquisition: A narrative review. *Applied Linguistics*, 36, 290–305.
- DeKeyser, R. M., Salaberry, M. R., Robinson, P., & Harrington, M. (2002). What gets processed in processing instruction? A response to Bill VanPatten’s “update.” *Language Learning*, 52, 805–823.
- DeKeyser, R. M., & Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning*, 46, 613–642.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–261). Cambridge: Cambridge University Press.
- Farley, A. P. (2005). *Structured input: Grammar instruction for the acquisition-oriented classroom*. New York: McGraw-Hill.
- Fernández, C. (2008). Reexamining the role of explicit information in processing instruction. *Studies in Second Language Acquisition*, 30, 277–305.
- Gregg, K. E. (1989). Second language acquisition theory: The case for a generative perspective. In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 15–40). Cambridge: Cambridge University Press.
- Henry, N. (2015). *Morpho-syntactic processing, cue interaction, and the effects of instruction: An investigation of processing instruction and the acquisition of case marking in L2 German* (Unpublished doctoral dissertation). Pennsylvania State University.
- Henry, N., Culman, H., & VanPatten, B. (2009). More on the effects of explicit information in instructed SLA: A partial replication of and a response to Fernández (2008). *Studies in Second Language Acquisition*, 31, 559–575.
- Lardiere, D. (2012). Linguistic approaches to second language morphosyntax. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 106–126). New York: Routledge.
- Lee, J. F., & Benati, A. (Eds.). (2013). *Individual differences and processing instruction*. London: Equinox.
- Lee, J. F., & VanPatten, B. (1995). *Making communicative language teaching happen*. New York: McGraw-Hill.

- Lee, J. F., & VanPatten, B. (2003). *Making communicative language teaching happen* (2nd ed.). New York: McGraw-Hill.
- Lightbown, P. M. (1983). Exploring relationships between developmental and instructional sequences in L2 acquisition. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 217–245). Rowley, MA: Newbury House.
- Marsden, E., & Chen, H. (2011). The roles of structured input activities in processing instruction and the kinds of knowledge they promote. *Language Learning*, 61, 1058–1098.
- Qin, J. (2008). The effect of processing instruction and dictogloss tasks on acquisition of the English passive voice. *Language Teaching Research*, 12, 61–82.
- Richards, J., & Rogers, T. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.
- Sanz, C., & Morgan-Short, K. (2004). Positive evidence vs. explicit rule presentation and explicit negative feedback: A computer assisted study. *Language Learning*, 54, 35–78.
- Sanz, C., & VanPatten, B. (1998). On input and output practice in second language acquisition: A response to Salaberry. *The Canadian Modern Language Review*, 54, 263–273.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Slabakova, R. (2012). L2 semantics. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 127–146). New York: Routledge.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14, 103–135.
- Uludag, O., & VanPatten, B. (2012). The comparative effects of processing instruction and dictogloss on the acquisition of the English passive by speakers of Turkish. *International Review of Applied Linguistics*, 50, 187–210.
- VanPatten, B. (1993). Grammar instruction for the acquisition-rich classroom. *Foreign Language Annals*, 26, 433–450.
- VanPatten, B. (1996). *Input processing and grammar instruction*. Norwood, NJ: Ablex.
- VanPatten, B. (2002). Processing the content of IP and PI research: A response to DeKeyser, Salaberry, Robinson and Harrington. *Language Learning*, 52, 825–831.
- VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Lawrence Erlbaum Associates.
- VanPatten, B. (2009). Processing matters in input enhancement. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 47–61). Bristol: Multilingual Matters.
- VanPatten, B. (2013). Mental representation and skill in instructed SLA. In J. Schwieter (Ed.), *Innovations in SLA, bilingualism, and cognition: Research and practice* (pp. 3–22). Amsterdam: John Benjamins.
- VanPatten, B. (2014). Input processing by novices: The nature of processing and research methods. In Z. Hong & R. Rast (Eds.), *First exposure to a second language: Learners' initial input processing* (pp. 193–207). Cambridge: Cambridge University Press.
- VanPatten, B. (2015a). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd ed., pp. 113–134). Mahwah, NJ: Lawrence Erlbaum Associates.
- VanPatten, B. (2015b). Foundations of processing instruction. *International Review of Applied Linguistics*, 53, 91–109.
- VanPatten, B. (2015c). Processing perspectives on pedagogical intervention. In J. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 200–215). Cambridge: Cambridge University Press.
- VanPatten, B. (2016a). *Language*. A mini-monograph for the e-book series *Foundations of language teaching*. New York: Routledge.
- VanPatten, B. (2016b). *Communication and skill*. A mini-monograph for the e-book series *Foundations of language teaching*. New York: Routledge.

Bill VanPatten

- VanPatten, B., Borst, S., Collopy, E., Qualin, A., & Price, J. (2013). Explicit information, grammatical sensitivity, and the first-noun strategy: A cross-linguistic study in processing instruction. *Modern Language Journal*, 97, 506–527.
- VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15, 225–243.
- VanPatten, B., Farmer, J. L., & Clardy, C. L. (2009). Processing instruction and meaning-based output instruction: A response to Keating and Farley (2008). *Hispania*, 92, 116–126.
- VanPatten, B., & Fernández, C. (2004). The long-term effects of processing instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 273–289). Mahwah, NJ: Lawrence Erlbaum.
- VanPatten, B., Inclezan, D., Salazar, H., & Farley, A. P. (2009). Processing instruction and dictogloss: A study on word order and object pronouns in Spanish. *Foreign Language Annals*, 42, 558–576.
- VanPatten, B., & Oikonen, S. (1996). Explicit information vs. structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495–510.
- VanPatten, B., & Rothman, J. (2014). Against “rules.” In A. Benati, C. Laval, & M. J. Arche (Eds.), *The grammar dimension in instructed second language acquisition: Theory, research, and practice* (pp. 15–35). London: Bloomsbury.
- VanPatten, B., & Rothman, J. (2015). What does current generative theory suggest about the explicit-implicit debate? In P. Rebuschat (Ed.), *Explicit and implicit learning of languages* (pp. 91–116). Amsterdam: John Benjamins.
- VanPatten, B., & Sanz, C. (1995). From input to output: Processing instruction and communicative tasks. In F. Eckman, D. Highland, P. W. Lee, J. Mileham, & R. R. Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 169–185). Mahwah, NJ: Lawrence Erlbaum.
- VanPatten, B., & Uludag, O. (2011). Transfer of training and processing instruction: From input to output. *System*, 39, 1–10.
- VanPatten, B., & Wong, W. (2004). Processing instruction vs. traditional instruction, once again: A study on the French causative. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 97–118). Mahwah, NJ: Lawrence Erlbaum.
- White, J. P., & DeMil, A. J. (2013). Primary and secondary effects of PI: A replication of Leiser and DeMil. *International Journal of Language Studies*, 7(3), 59–88.
- White, J. P., & DeMil, A. J. (2014). Frequency effects of structured input: Primary and secondary effects. *Southern Journal of Linguistics*, 38(2), 1–46.
- White, J. P., DeMil, A. J., & Rice, M. (2015). Traditional instruction and processing instruction: German dative definite articles. *German as a Foreign Language*, 15(1), 1–26.
- White, L. (2015). Linguistic theory, Universal Grammar, and second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd ed., pp. 34–53). New York: Routledge.
- Wong, W. (2004). The nature of processing instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 33–63). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wong, W. (2005). *Input enhancement: From theory and research to the classroom*. New York: McGraw-Hill.

Assessment in the L2 Classroom

Ute Knoch and Susy Macqueen

Background

Assessment processes are present in all second language (L2) classrooms, regardless of the approach to instruction. They range from intuitive, moment-by-moment teacher decisions and responses on the one hand, to the formal delivery of tests and scores on the other (Genesee & Upshur, 1996; Leung, 2005; Rea-Dickins, 2001). A key feature of classroom-based assessment (CBA) is that it is qualitatively closely connected to, or even embedded in the activity of learning, and as a result, CBA focuses primarily on the trajectories of individual learners or very small groups. At the same time, CBA is usually also connected to externally imposed expectations about instruction and learning that relate to the trajectories of larger groups and populations (see also discussion in Black & Wiliam, 1998). Hence, classroom assessments are typically constructed in relation to some kind of external standard, whether it is a self-contained external instrument such as a standardized test, or a pedagogical/developmental map such as a syllabus or curriculum framework and its associated outcomes. In relation to such standards, teachers are required to assign grades to a temporal collection (e.g., a course, a semester) of classroom-embedded assessments, which frequently involves distinguishing between assessments that are primarily learning experiences (termed ‘formative’ assessment or assessment *for* learning) and those which are primarily informative about the ultimate gain from the period or type of learning (‘summative’ assessments or assessment *of* learning) (Black, Harrison, Lee, Marshall, & Wiliam, 2003; Wiliam, 2011). In this chapter, we present various types of CBA. We discuss how different approaches and methods might link to individual learner trajectories as well as to external standards. We also explore the kinds of decisions teachers must make in relation to assessment to ensure that what, when, how, and why they assess is congruent with their approach to instruction, the individual learners and external standards.

Classroom-Based Assessment

Classroom-based assessment (CBA) has a very broad remit. Leung (2005) construes it as “noticing and gathering” information during ordinary classroom activities for decisions about teaching without necessarily quantifying or reporting (p. 871). Jamieson (2011)

Ute Knoch and Susy Macqueen

contrasts this daily decision-making with more formal testing processes that are explicitly linked to a classroom-external *domain of use*. The difference between day-to-day formative assessing (e.g., incidental feedback) and mandated summative assessments (e.g., a standardized test) is the extent and explicitness of the connection between the assessment and classroom-external expectations, such as a state-prescribed curriculum framework and benchmark performances. In the case of teachers' noticing of learners' performance during classroom activities, the judgement is very local; it is embedded in the choice of activity and what it elicits from individual learners. However, such embedded assessment processes may well be affected by classroom-external expectations; consciously or otherwise, classroom-internal assessment practices are influenced or even dictated by predominant large-scale testing regimes, national curricula and standards, and their inherent views of language and language learning in interaction with teacher beliefs (Spratt, 2005).

Formative assessment, then, refers to a range of continuous evaluative processes carried out as part of the teaching and learning processes, with the potential to inform and improve both (see, for example Wiliam, 2011). Formative processes can include peer- and self-assessment and qualitative feedback of various types. As Bachman (1990) points out, such decision-making is relatively low-stakes and potentially reversible. In contrast, summative assessment carries higher stakes because it usually involves formal grading and reporting. Hill and McNamara (2012) propose a useful framework that encompasses both formative and summative processes under the assessment dimensions of *evidence*, *interpretation*, and *use* (McNamara, 2001). Hill and McNamara construe these dimensions as a series of questions or issues, shown in Table 11.1. In this chapter, we use two examples of assessment to illustrate and discuss teacher decisions about assessment design, implementation, and purpose as set out in Hill and McNamara's (2012) framework. The two examples are:

Example 1

Context: A 10-week academic English language pathway course—final course grades determine whether or not students can enter university degree courses in Australia.

Task: An academic literature review assignment on a topic related to students' future university disciplines. Assignments are completed in groups of three students and the collective final grade contributes 20% to the overall final grade for the course. The criteria used to determine assignment grade are: (1) grammatical range and accuracy, (2) lexical resource, (3) structure and argumentation, and (4) referencing.

Example 2

Context: A general English language class at a private language school in Japan.

Task: The students carry out a writing activity about what they did on the weekend, which elicits, among other structures, the use of the past simple tense. The teacher gives corrective feedback on grammar and vocabulary.

Table 11.1 Dimensions of CBA with examples

<i>Dimensions</i>	<i>Teacher considerations</i>	<i>Operationalization</i>	<i>Example 1</i>	<i>Example 2</i>
EVIDENCE	What is assessed?	Prioritized skills and aspects of language	<i>Academic literature review on a discipline-specific topic</i>	<i>The use of the past simple tense in informal writing</i>
	How is evidence collected?	Assessment event may be (1) planned or incidental (2) visible to learners or embedded in teaching	<i>Group assignment with peer and teacher feedback on first draft; following revisions, teacher gives final grade and further feedback</i>	<i>In-class writing activity; teacher provides feedback on grammar and vocabulary by identifying errors for students to fix</i>
	When?	(1) Beginning of term (2) End of term (3) Regularly throughout term	<ul style="list-style-type: none"> <i>Week 4: feedback on draft (peers & teacher)</i> <i>Week 6: final grade (teacher)</i> 	<i>As part of a regular writing activity in first lesson of the week</i>
	Who is assessed?	(1) Individual (2) Group (3) Class	<i>Group collaboration; same mark for each group member</i>	<i>Individual learner</i>
	By whom?	(1) Teacher (2) Learner	<i>Peers, Teacher</i>	<i>Teacher</i>
INTERPRETATION	What is the level of attention by the teacher and learners?	(1) Sustained (2) Momentary	<i>Sustained attention over time</i>	<i>Momentary</i>
	What are the criteria/values/standards guiding the assessment?	(1) Explicit or unconscious (2) External or indigenous	<ul style="list-style-type: none"> <i>Criteria show what teacher/institution believe to be relevant and worthwhile domain (university) practices</i> <i>Teacher/peer feedback contribute to learning</i> 	<ul style="list-style-type: none"> <i>Grammatical accuracy is necessary for language proficiency (equal to fluency)</i> <i>Corrective feedback is an effective practice</i>
USE	How is evidence used?	<ul style="list-style-type: none"> <i>Assign level</i> <i>Plan/modify teaching</i> <i>Learning</i> <i>Management</i> <i>Socialization into culture of assessment</i> 	<ul style="list-style-type: none"> <i>Assignment grade contributes 20% to overall grade, which determines entry to university degree courses</i> <i>Students may apply feedback to subsequent academic writing</i> 	<ul style="list-style-type: none"> <i>Teacher locates error and student may fix it, but no follow up by teacher</i> <i>Teacher may focus on grammatical feature in future instruction if error observed widely</i>
	By whom?	<ul style="list-style-type: none"> <i>Teacher/learner</i> <i>School</i> 	<i>Admitting institution and student</i>	<i>Teacher and student</i>

Source: Adapted from Hill & McNamara, 2012.

As can be seen in Table 11.1, a teacher's evaluation of an individual's language learning is determined via an amalgam of decisions, listed under 'teacher considerations.' Each of these decisions is underpinned by the general approach taken to instruction. Therefore, instructional approaches that prioritize scaffolded performance might assess in the vein of the first example, which includes collaborative group work. A teacher whose approach prioritizes a focus on language structure might formatively assess along the lines of the second example where the feedback is focussed on grammar and vocabulary. We have added timing (*When?*) as a consideration to Hill and McNamara's original list because of its significance in different theoretical frameworks, for example the time it takes to develop self-regulation through collaborative interactions with peers or expert–novice interactions in socioculturally oriented approaches (Lantolf, Thorne, & Poehner, 2015) or an immediate, incidental response in a meaning negotiation that has been triggered by a learner's language communication needs in form-focused approaches (Long & Robinson, 1998). Regardless of what approach is taken to instruction, it can be seen that assessment processes are inherent in instructed second language acquisition (ISLA), a fact that makes it paramount that teachers engage in informed and principled decision-making with regard to their assessment instruments and processes.

In addition to illustrating pedagogical approaches, the two examples in Table 11.1 also show how classroom assessment processes vary in terms of how externally connected and accountable they are. The first example, a rather high-stakes assignment that contributes to a final grade that is used to determine university entrance, appears to be domain-referenced, that is, it is intended to reflect the genre of the target domain (in this case, a literature review) and domain-relevant assessment criteria (e.g., appropriate referencing is considered to be a key skill in the target university domain). It has both formative and summative aspects (formative—peer and teacher feedback on drafts to promote individual development; summative—an overall contribution to a high-stakes use in relation to a population of university entrants). The second example, an instance of corrective feedback during an informal personal writing task, is embedded in a learning activity and has no apparent official use other than to promote language development. Despite its individualized, momentary nature, the use of corrective feedback may well be due to external measures such as a grammatical criterion in a relevant standardized test or the teacher's experience and beliefs about the nature of language and language learning (see Hill & McNamara, 2012). Further, it is quite possible for contradictions and tensions to exist between the external expectations imposed on assessment practices (for example, institutional policies) and the teacher's preferred approach to instruction (documented in Alderson & Hamp-Lyons, 1996). For example, a teacher may be opposed to grading formative assessment (i.e., using formative assessment for summative purposes), but this may occasionally be required as part of the institution's policy. We would argue that teachers should be aware of the connections between their learner-oriented assessments of individuals and the external prescriptions or influences on these. CBA provides a critical nexus among (1) the characteristics and needs of individual students (e.g., motivation, time available, personality, prior language learning experience); (2) the affordances of the classroom experience, including but not limited to the instructional approach/es adopted by the teacher; and (3) classroom-external expectations of what the instructional period should produce.

The final point to note about CBA is that the type of evidence gathered is varied. Genesee and Upshur (1996) describe a broad range of assessment methods that produce different types of evidence and contribute in various ways to the compilation of an overall picture of a learner's progress and/or general language ability. These methods include observations,

portfolios, conferences, journals, questionnaires, interviews, test-tasks, rating scales, and checklists. Before we deal in some depth with selected key methods of gathering evidence about individuals' language learning in classrooms, we will discuss the issue of external standards or proficiency frameworks and how these relate to CBA and ISLA.

Key Concepts

Classroom-based assessment: Frequently contrasted with standardized testing, classroom-based assessment is carried out as part of learning and teaching for the purposes of informing future teaching.

Formative assessment: Assessment that is primarily a learning experience, aimed at determining the most appropriate future learning experience(s).

Summative assessment: Assessment that is designed to be informative about the ultimate gain from the period or type of learning.

Teaching Tips

- Make yourself aware of the external forces (e.g., tests, curriculum outcomes) that exert influence on your assessment practices. Ask yourself:
Why am I assessing [this skill/genre/language point] in this way?
What decisions or actions does this assessment contribute to?
- Be prepared to act on your reflections. For example, you might consider changing an assessment task if it is difficult to justify in relation to (1) the domain or external test of relevance to the course or (2) the students' needs.
- Give students a clear rationale for the assessment tasks you set so they are aware that you see assessment as part of their learning and your overall teaching plan, rather than something separate.

The Role of Standards in CBA

Most classroom teachers are required to report learner performance against a set of external standards, such as proficiency scales or frameworks (e.g., the Common European Framework of Reference), which are designed to provide a common set of descriptions against which curricula and tests can be benchmarked. External standards are generally implemented because of the need for accountability as governments, educational boards or education providers need to ensure that students achieve certain demonstrable outcomes from their education and that these outcomes are documented in a uniform manner. Standards are generally formulated as a series of statements in terms of 'stages' of language behaviour ranging along a continuum from 'close to zero' to 'near native-like' and are designed to represent different skills and typically describe different texts and tasks learners can handle at a number of different levels (Brindley, 1998). When these kinds of frameworks are applied across institutions, the external measure is also a means of introducing marketplace competition into the education sector so that stakeholders can make informed choices (McKay, 2000).

Besides offering accountability and uniformity across school systems or sectors, external frameworks or standards can also bring about fairer assessment practices. In an English-speaking school system, for instance, groups of learners who might be categorized as very low achievers against native English-speaking peers can be recognized as bilingual learners who are attaining predictably in relation to bilingual/bicultural developmental expectations as charted in an appropriate framework (for example, *Bandscales for Aboriginal and Torres Strait Islander Learners*, Angelo, 2013). These kinds of instruments can also have a professional development function for teachers who have not undergone L2 teacher training, but who find themselves responsible for assessing L2 learners (Macqueen, Harding, & Elder, 2011).

However, frameworks or standards¹ have been criticized for a number of reasons. First, it is questionable how well these scales align with findings from SLA research and represent actual developmental sequences or learning trajectories at the level of individual learners (e.g., Fulcher, 2003). SLA studies investigating developmental sequences for different aspects of language have often found evidence of nonlinear development (e.g., Meisel, Clahsen, & Pienemann, 1981), which means that learners do not progress in linear fashion but may show a U-shaped or N-shaped developmental curve, which is not captured by such frameworks. While this criticism is certainly valid, it is important to point out that developmental hierarchies in SLA generally take a fairly narrow focus (e.g., on morphosyntax) while frameworks usually take a broader or more holistic view of learners' proficiency. A related criticism to the one previously mentioned has been that proficiency frameworks usually also lack a theoretical basis as they are often not based on a theory of language ability or development (see e.g., Bachman, 1990; Bachman & Palmer, 1996) but on intuitions of test developers or teachers. Scales have also been criticized for a lack of generalizability as the ability levels are usually characterized in terms of what learners 'can do' on specific texts and tasks. These texts and tasks usually change substantially across proficiency levels. Doubts have also been cast about the validity of the hierarchy of text and task types in such frameworks (see for example Lee & Musumeci, 1988 who list different texts and tasks a learner is able to produce or understand at different developmental levels).

Curriculum standards are powerful policy documents, as they specify expected language development outcomes that teachers then base their assessment tasks (and scoring/feedback) upon. Therefore, curriculum standards represent the construct of the assessment, that is, they are a representation of what CBA should assess in the view of the government body. As McNamara and Elder (2010) argue, in standards-based systems assessments are designed to investigate what the standards include, and nothing else. In this way, the standards restrict what can be measured and reported, and as a direct result constrain the possible constructs that may be addressed in a classroom. This restriction is particularly problematic as it is difficult to change standards because they are often policy documents implemented at high institutional and governmental levels.

Frameworks or standards are reductionist not only in terms of the constructs included, but also in levels. Standards often contain a limited number of levels, for example 6 or even 10 levels, progressing from total beginners to highly proficient users of language. While this may sound sufficient, this number of levels is often very limiting for classroom teachers, who would like progress to be described on a much finer level to show smaller increments of progress made by learners and to avoid frustration on behalf of students and parents. This is a clear shortcoming of external standards in the context of CBA, in particular in the case of assessment for learning, that is, assessments designed to provide regular feedback on learning progress.

Classroom teachers have the added challenge of ‘linking’ or basing their classroom language assessments on standards or frameworks that are often very general in nature and provide little guidance for the purpose of assessment development practices. A useful notion in this regard is that of the ‘assessment cycle’ (Rea-Dickins, 2001), which comprises a sequence of four decision-making stages (p. 435). The stages encompass both learning-oriented classroom processes and their interpretation in relation to institutional and external requirements. The initial planning stage involves the teacher identifying the purpose, the method, and the agents and preparing the learners for the assessment. The second stage is implementation and includes processes such as scaffolding, feedback, and monitoring. The third stage is the beginning of interpretation of the evidence gathered about each learner. At this point, further feedback (broadly conceived) is provided to learners and instructional plans may be revised. The final stage involves formal review for internal school purposes and recording progress against standards.

Key Concepts

External standards: A framework that describes levels of attainment, often in a series of learning outcomes that are used for reporting purposes (as well as throughout the teaching cycle). Such standards may also be expressed in a single testing regime where scores are arranged in descriptive levels that are matched to course levels or outcomes.

Assessment cycle: The classroom and institutional stages in assessing students, including planning, implementing, interpreting, and reporting.

Current Issues and Empirical Evidence

The remainder of this chapter focuses on classroom assessments that fit with an assessment cycle involving some degree of planning and emphasizing learning as a joint enterprise between the teacher and the individual learner through negotiated processes such as feedback and scaffolding, but are nonetheless accountable to external expectations such as curriculum standards. We consider research that examines the efficacy of these approaches in terms of language learning and acquisition, and we relate this research to Hill and McNamara’s (2012) model of classroom assessment.

Teacher Written Corrective Feedback

Teacher written corrective feedback can be viewed as a type of language assessment, as it involves the teacher reviewing a performance and providing feedback to a learner. Both teachers and learners gain an understanding of error patterns, and that knowledge can be used for future teaching and learning. Written corrective feedback, therefore, is a type of language assessment that is very much intertwined with teaching processes. The role of written corrective feedback to student writing has increasingly been studied in recent years, in particular because a number of inconclusive early studies investigating the effectiveness of such feedback triggered Truscott (1996) to take a strong position against the provision of written corrective feedback. Truscott argued, among other things, that written corrective feedback at best benefits explicit knowledge about language and not implicit knowledge,

which is the knowledge employed when language is used (DeKeyser, 2003). Truscott therefore contended that this kind of feedback can merely prompt pseudo-learning and may only help students to improve their self-editing skills when writing. Truscott's arguments triggered a large body of research on written corrective feedback, described further next, which has largely refuted his line of reasoning.

Written corrective feedback has been classified in a number of ways, which are worth exploring further before turning our focus to the research done in this area. A distinction has been made between focussed and unfocussed feedback (Bitchener & Ferris, 2012). When providing the former, the teacher or researcher provides feedback on one or a limited number of structures only, disregarding any other errors. In the latter form of feedback, teachers provide feedback to a large number, if not all errors they encounter. Unfocussed written corrective feedback is arguably the most common type practiced in the classroom, although teachers invariably differ in the extent of such feedback (see, for example Alshahrani & Storch, 2014; Guenette & Lyster, 2013). Some teachers may decide to focus only on structures that have already been taught in the curriculum; others may focus on aspects of language seemingly 'within reach' for a particular learner; still others may provide feedback to all errors they notice. It is clear from this description that unfocussed feedback may be far from standardized across learners and classrooms.

A further distinction has been made between direct and indirect feedback. Direct feedback refers to the provision of the correct form. Indirect feedback practices can vary but involve an indication that an error has occurred without supplying the correct form (Ferris, 2011). Indirect feedback can be provided through underlining or circling the error, or through the use of error correction symbols or codes. It is then up to the learner to correct the error.

Before reporting findings of the efficacy of these different types of written corrective feedback, it is important to mention that methodological designs of some studies have made it difficult to draw firm conclusions. For example, some studies failed to measure the effectiveness of the feedback on new pieces of writing, requiring students only to revise their essays. Effectiveness is best measured by showing that students are able to improve their writing in completely new pieces as revisions may merely indicate successful copying of direct feedback. Other studies did not include a control group, making it difficult to conclude whether any improvement was truly because of the feedback provided. Finally, many studies employed a one-shot design, providing feedback only once rather than several times as would be the case in many language classrooms. Studies in which feedback is only provided once may not be able to show improvement because students may need to notice their errors more than once. For this reason, proponents of corrective feedback (e.g., Bitchener & Ferris, 2012) have argued that feedback in research studies should be provided more than once to students, adding ecological validity to these studies (Storch, 2010). Studies providing repeated feedback over, for example, a semester, have been able to show that this method is effective (see e.g., Rastgou, 2016).

Despite the shortcomings of some, mostly earlier studies, research findings are increasingly providing a picture of the effectiveness of written corrective feedback. For example, studies examining focussed written corrective feedback (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009a, 2009b, 2010a, 2010b; Bitchener, Young, & Cameron, 2005; Sheen, 2007) have all shown that such feedback is able to improve accuracy in new pieces of writing completed after a period of time. Studies comparing focussed and unfocussed approaches to the provision of written corrective feedback (Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, Wright, & Modawa, 2009) have been mixed, with one study resulting in accuracy gains only for the focussed feedback group (Sheen, Wright, & Modawa, 2009) and the other

showing increases in accuracy in both groups (Ellis, Sheen, Murakami, & Takashima, 2008). A recent study by Van Beuningen, de Jong, and Kuiken (2012), comparing the effects of direct and indirect unfocused written corrective feedback, showed that both methods were effective in reducing errors in both new and revised pieces of writing. Indirect written corrective feedback was shown to be most effective in reducing nongrammatical errors (e.g., those related to appropriateness, vocabulary, and spelling).

Research on feedback provision beyond one-shot designs is less common. In a recent study (Rastgou, 2016), feedback was provided eight times (on a large range of preselected structures) in a teaching term followed by a posttest and a delayed posttest 4 weeks later. The study showed that students' writing improved in accuracy and that this was sustained over time.

Written corrective feedback therefore has been shown to be effective in an increasingly large number of studies, providing teachers with evidence that their practices, if well delivered, can help students improve their accuracy in writing. However, providing feedback to learners is time-consuming, and for this reason peer assessment, which will be discussed later in this chapter, is an attractive alternative.

Diagnostic Assessment and Feedback in the Language Classroom

Diagnostic assessments are designed to provide detailed feedback about learner's strengths and weaknesses (Alderson, 2005). For that reason, diagnostic assessments focus on more specific abilities than, for example, proficiency tests and placement tests and are quite difficult to develop. Classroom-based diagnostic assessments are usually directly related to the course syllabus and may be administered at the beginning of a unit to provide the teacher with information about the level of the learners in relation to the upcoming teaching material or after a unit of instruction to give detailed feedback to the learners and teachers about what aspects of language have been learned and what aspects need further attention. In this way, teachers can use the information learned from the assessment for planning their upcoming instruction and learners can use the information to guide their own learning. To enable detailed feedback, teachers or test developers need to clearly define what subskills they would like to assess and provide feedback on when they are developing the assessment. For diagnostic assessments of receptive skills such as reading and listening, this is a particularly difficult task as it is often difficult for a group of teachers from the same context to fully agree on what every item is designed to assess. Studies such as Lumley (1993) have shown that experts (e.g., test developers or teachers) disagree with each other on what subskills particular test items measure. Some test questions/items are also commonly identified to be measuring more than one particular subskill (see e.g., Harding, Alderson, & Brunfaut, 2015). Because a subskill analysis is critical in providing detailed feedback on strengths and weaknesses in a learner's knowledge and use of language, this issue raises a potential threat to the validity of the assessment. It may be easier for classroom teachers to develop diagnostic assessments of grammar or vocabulary knowledge, basing their test items on taxonomies of grammar skills (Purpura, 2013), for example. Diagnostic assessments of writing and speaking are more commonly employed in language classrooms.

To provide optimal feedback, teachers need to decide on the level of detail of the feedback provided to learners. For example, diagnostic feedback following a paragraph writing activity could either tell the writer about the accuracy of the grammatical structures used or it could give detailed feedback on a number of grammatical structures, in particular those that were recently covered in the course book. This choice is usually directly related to the

material covered in an upcoming or previous unit of instruction. Feedback is often presented in graphical format (by showing progress toward mastery on graphs representing different subskills), if possible, to enable learners to better understand the results. The role of such feedback is not to merely provide learners with error correction, but to address the gap between a current level of achievement and a desired level (e.g., Jang & Wagner, 2014). To make this gap more apparent, diagnostic assessment is often administered in combination with self-assessment (Oscarson, 2014). In this way, diagnostic assessment is designed to make learners active participants in the learning process by gaining a deep knowledge of the criteria and their position on a developmental continuum.

Research on classroom-based diagnostic language assessment is surprisingly rare (see, for example, Jang & Wagner, 2014; Knoch, 2009). Little is known about how learners engage with the feedback (but see Jang, Dunlop, Park, & van der Boom, 2015), how teachers use the results of such assessments to feed into their future teaching, and most importantly of all, whether the results have an impact on learning and acquisition of language. While there is evidence of the effectiveness of diagnostic approaches to first language acquisition (see, for example, Alderson, 2005; Huhta, 2008), such studies are clearly needed to establish the possible effects of providing diagnostic feedback for L2 learning and retention. Knoch (2015) introduced the concept of developmental diagnostic writing assessment, arguing that writing knowledge is built over time and across a number of genres and that developmental diagnostic feedback would provide both teachers and learners, who generally spend at least one term or semester together, detailed evidence of growth over a period of time. Whether the concept of developmental diagnostic assessment is beneficial to learning needs to be empirically tested.

Peer Assessment

Peer assessment refers to “an arrangement of peers to consider the level, value, worth, quality, or successfulness of the products or outcomes of learning of others of similar status” (Topping, Smith, Swanson, & Elliott, 2000, p. 150). Peer assessment can focus on both written and oral skills and can be implemented in pairs or groups. Students can provide feedback on a range of task types, including writing assignments, portfolios, and oral presentations, although it is most commonly used in the L2 classroom for writing (Hansen Edwards, 2014). Peer feedback has a number of benefits apart from saving teachers time. It requires higher order thinking processes from both the reviewers and the feedback receivers, including the ability to develop problem-solving skills. In the process, students develop ownership of the assessment process and gain a much deeper understanding of the assessment criteria. By using a combination of teacher and peer feedback in classes, students receive a greater quantity of feedback and possibly also faster feedback. Through the social interaction, learners are actively involved in the learning process, and gain independence from the teacher.

There are also a number of potential drawbacks to peer assessment. To be effective, peer feedback requires class time for training learners in the optimal use of peer feedback techniques that may be time-consuming. Depending on their cultural background, students may be unwilling to assess their peers or they may prefer teacher feedback. Most importantly, however, students may not have the linguistic knowledge to comment on the accuracy of writing and as a result they may provide incorrect feedback to their peers (Hansen Edwards, 2014).

While the use of peer feedback is intuitively appealing, it is important to examine its effectiveness empirically. A number of studies have compared peer and teacher feedback.

The findings show that learners are more likely to incorporate feedback from teachers than from peers (e.g., Paulus, 1999). However, peer comments were useful in some areas, for example when raising writers' awareness of audience (see, for example, Tsui & Ng, 2000), suggesting the complementary value of peer and teacher feedback. Similarly, studies by Yang, Badger, and Yu (2006) and Zhao (2010) have shown that peer feedback can result in a higher proportion of meaning-level revisions as opposed to the surface-level revisions common following teacher feedback. These studies also showed that learners seem to understand a greater proportion of peer comments because the language used by teachers may be confusing or difficult to understand.

A number of factors affecting the efficacy of peer feedback have also been investigated. These studies have shown that the nature of the interaction in the peer dyad can influence the effectiveness of peer feedback (e.g., de Guerrero & Villamil, 1994, 2000; Nelson & Murphy, 1993). Students have been shown to engage in a number of different interaction patterns, including cooperative and defensive. Cooperative dyads have been shown to incorporate more of the comments. L2 proficiency level has also been shown to be a factor, with high proficiency learners benefiting more from such feedback than their lower proficiency counterparts (see, for example Kamimura, 2006). Finally, training in peer assessment has proven to influence the effectiveness of peer feedback. For example, Rahimi (2013) was able to show that training resulted not only in higher quality comments among the peers, but also in more accurate writing performance in new pieces of writing.

While the studies examining the efficacy of peer feedback have varied in focus (e.g., evaluating the quality of the comments provided by peers, or comparing the peer comments to teacher feedback or self-assessments) only one study (Rouhi & Azizian, 2013) that we are aware of has implemented the type of design advocated by researchers in the area of written corrective feedback who argue that it is not enough to show increased quality in revised written samples, but that this improvement needs to be sustained in delayed new pieces of writing. Rouhi and Azizian (2013) examined improvements in new pieces of writing, but focussed only on improvement of the writing in feedback givers, as this was the focus of their study. For this reason, it is difficult to make any firm claims about the value of peer feedback on the automatization of grammatical structures or other aspects of writing for which learners have prior explicit knowledge. Further studies need to establish whether short-term gains found in many studies can be retained by learners.

Dynamic Assessment

Dynamic assessment of language is a form of CBA that is, just like written corrective feedback, closely connected to teaching and learning. In fact, Poehner and Infante (2016) argue that in the case of dynamic assessment, teaching, and assessment are understood as "dialectically related features of the same educational activity" (p. 1), which is designed to promote student learning. While most forms of assessment are designed to collect data on a learner's proficiency or ability to use language, the aim of dynamic assessment is to provoke learner development during the activity. Similarly, while in other types of assessment no help from external sources like teachers or dictionaries is provided to the learner, dynamic assessment relies on the provision of different levels of scaffolding to the learner to examine how responsive learners are to the provision of such help. The level of support or scaffolding needed to complete an item provides the teacher with information about how far from independent functioning a learner is. The support offered during a dynamic assessment activity is at the same time intended to promote language development.

The tenets of dynamic assessment derive from sociocultural theory (Vygotsky, 1978) and the idea that development occurs through mediation (Lantolf & Poehner, 2014). With a typical language assessment, two students may receive the same score. However, with dynamic assessment one learner may answer more items correctly with the help of mediation (either through interaction with a teacher or through other resources), while the other may not benefit from such help. The first learner, therefore, is closer to independent mastery of these items and therefore at a higher developmental level than the second learner. More traditional types of language assessments would not show the difference between the two learners while these differences can be drawn out by dynamic assessment. Vygotsky (1978) therefore argues that conventional assessments can only provide evidence of learners' 'zone of actual development' rather than the 'zone of proximal development.'

A number of different approaches to dynamic assessment have been described. In interactionist dynamic assessment approaches (Lantolf & Poehner, 2004), learners interact directly with a teacher who provides mediation in the form of different levels of scaffolding as learners encounter problems with answering an item. Because these interactions are open-ended and not scripted, the results of the assessment are qualitative profiles of learners, which may be difficult to compare across different students. For this reason, interactionist approaches to dynamic assessment are not always suitable beyond classroom contexts. Interventionist approaches, on the other hand, standardize the mediation offered as part of the assessment (Poehner & Infante, 2016). More standardized mediation in dynamic assessment can be achieved by scripting the interaction so that it is offered in a standardized format (usually from least to most explicit) or through computer programs with preprogrammed mediation (e.g., Poehner & Lantolf, 2013; Poehner, Zhang, & Lu, 2015). Poehner (2009) also proposed the idea of group dynamic assessment in which groups of students together work on tasks slightly beyond their reach but achievable with mediation.

Outcomes of dynamic assessments can provide educators with much more information than the raw scores of conventional assessments. Poehner et al. (2015), for example, experimented with an actual score (which represents independent performance), a mediated score (which indicates how many mediational 'levels' a learner required across the test), a transfer score (which shows whether learners were able to transfer ideas 'learned' through mediation to new, more difficult items), and a learning potential score (which indicates the degree of progress made by learners during the activity) (see also Kozulin & Garb, 2002). Poehner et al. (2015) were able to show that students with the same raw scores received vastly different score profiles across the other areas, showing that students benefitted very differently from mediation and that different amounts of learning resulting from mediation could be transferred to similar, new items. No dynamic assessment studies to date, that we are aware of, have examined whether acquisition takes place from taking part in a dynamic assessment. Transfer scores provide some indication in that direction, i.e., how much of the learned material can be applied to new tasks. While this transfer to new tasks is not directly a sign of acquisition, this line of research is certainly an avenue that needs to be explored further. Poehner et al.'s (2015) study is also one of the first that has examined the diagnostic potential of dynamic assessments by providing information on subconstructs tested. They provide the example of two learners who received exactly the same actual scores, mediated scores, and transfer scores, but showed strengths and weaknesses in different areas of language (e.g., vocabulary and grammar), which indicates that there may be potential to merge the two areas of dynamic and diagnostic assessment in future work.

Portfolio Assessment

Most CBA is based on collecting and evaluating individual performance samples. However, teaching is inherently concerned with a bigger developmental picture than is shown in ‘one-shot’ views. Similarly, generalizing performance from one performance to a universe of performances (Weigle, 2002) is problematic. Portfolio assessment is a more developmental approach where students are assessed progressively through a collection of work samples done throughout a period of instruction (Davies & Le Mahieu, 2003). This process allows a more ‘time-lapse’ view of language development, where learners might, for example, work on several drafts of a piece of writing under the guidance of a teacher, produce a particular genre more than once in a teaching period, collect samples of various genres and skills over the period of instruction or some combination of these. A portfolio can be defined as a purposeful collection of student work collected over a period of time (Danielson & Abrutyn, 1997; Weigle, 2002). In L2 learning, the main adoption of this method has been in the field of L2 writing, although portfolios can encompass all aspects of language and various different assessment methods.

The use of portfolios can be mainly formative where the final evaluation is delayed so that learners have the opportunity to revise products before submission (Hamp-Lyons & Condon, 2000). When portfolios are used summatively (e.g., for reporting purposes), critical components from the entire period of instruction are represented. A more summative use of portfolios ideally forms an overall feedback process that can be used in future periods of instruction and may well be used in lieu of more formal testing procedures. The sampling process is more ecologically valid than sampling under test conditions, as learners can produce language samples in ways that are less affected by test methods and arguably closer to the writing practices of many domains.

Although portfolio compilation allows a more comprehensive, varied, and flexible sampling of student work, the nature of the compilation needs to be appropriate to the use of the resulting scores (Davies & LeMahieu, 2003), e.g., a portfolio that includes a scaffolded drafting process and a final polished work would not provide evidence that a learner is able to manage in a domain that requires control of several genres. Principles also need to be adopted in terms of what is given greatest emphasis in the final evaluation. Considerations such as how to balance the representation of assisted and independent work have to be weighed in relation to the purposes of the overall portfolio grade. In one type of implementation of portfolio assessment, for instance, although the entire portfolio is considered, the evaluation places weight toward the “fullest and latest” samples that are more recent and that provide evidence on all mandatory aspects of the curriculum (Maxwell, 2004, p. 4). Furthermore, with more comprehensive sampling comes a need for time-intensive moderation processes where teachers must take into consideration a potentially quite complex array of tasks and samples to ensure consistent summative grading.

Portfolio assessment therefore can include elements of all the assessment methods we have described so far; it enables different types of evidence to be compiled in a principled manner. It can contain assessment, which is more expert-mediated, as in dynamic assessment, as well as static samples. An important feature of portfolio assessment is that it allows learners some agency in the assessment process (Danielson & Abrutyn, 1997). Students may, for example, reflect on their own development as a part of the sample collection process or select which work is included in the portfolio. In some uses, the students are encouraged to articulate reasons for sample inclusion using the language of the task criteria (Davies & LeMahieu, 2003), which in turn reflect the standards. This kind of use emphasizes the importance

Ute Knoch and Susy Macqueen

of articulating criteria in ways that are useful for learners, teachers, and accountability purposes.

Pedagogical Implications

The assessment methods and processes presented earlier are ones that we consider to have the greatest level of interaction with approaches to ISLA. Mostly they are classroom-embedded methods that focus on the individual trajectories of learners and feed primarily and directly into teaching and learning. Although all types of assessment reviewed earlier (teacher written corrective feedback, diagnostic assessment, peer assessment, dynamic assessment, and portfolio assessment) are pedagogically motivated, it is useful to view them in terms of the kinds of evidence they provide, the interpretation of this evidence and the information they provide beyond the classroom context. For this reason, we have set them out in relation to teacher decision-making considerations as expressed in Hill and McNamara's (2012) framework. These are shown in Table 11.2.

For most teachers, the 'what is assessed' consideration in the first row of Table 11.2 is relatively straightforward. Awareness of the use and purpose of the assessment (bottom rows), should form the basis of other decisions about how to collect evidence, when to assess, and so forth. This then involves being aware of the values underpinning the assessment, including what theory or view of SLA is represented in the standards and course outcomes (Hill & McNamara, 2012). Diagnostic assessment of listening skills, for example, ideally relies on a clearly articulated model of L2 listening (Harding et al., 2015). Dynamic assessment is based on an understanding that the level of self-regulation is a critical consideration in promoting L2 learning (Aljaafreh & Lantolf, 1994). These kinds of understandings are not necessarily congruent with the broad articulations of language development commonly seen in attainment standards and it is therefore important that teachers are made aware of these issues.

Clearly, these considerations interact also with the approach to instruction adopted. Whether or not a teaching approach prioritizes form or meaning (as set out in Loewen, 2011), for example, has implications for the timing and nature of the assessment method. If the teaching syllabus is explicitly forms focussed, a grammatical diagnostic assessment at the beginning of the course would be congruent with the teaching approach. If the approach is meaning-focussed, incidental attention to form might occur through providing written corrective feedback that corresponds with particular communication needs. Conversely, a task-based pedagogy, which is more meaning-focussed, may not easily lend itself to the discrete information provided in diagnostic assessments (Harding et al., 2015). While it is worth being aware of these kinds of mismatches, teachers are likely to be much more flexible in their approaches than envisaged in theoretical discussion, and assessment methods may well be institutionally prescribed but teaching methods left up to the individual teacher.

A related consideration is the level of attention given to any assessment as a whole or aspects of it. Attention and noticing are central in many theories of ISLA, but when considered in relation to assessment methods, it is important to see attention as a more distributed notion that is realized, to some extent, as a result of assessment practices. Gass's (1997) computational model, for example, puts noticing as an important stage in moving from input processing to modified output. She argues for any feedback to be effective, learners need to notice the gap between their current interlanguage and a correct form. Different language assessments in the classroom would result in different levels of attention by the

Table 11.2 Teacher considerations relating to different assessment methods based on Hill and McNamara's (2012) dimensions of CBA

<i>DIMENSIONS</i>	<i>Teacher considerations</i>	<i>Written corrective feedback</i>	<i>Diagnostic assessment</i>	<i>Peer assessment</i>	<i>Dynamic assessment</i>	<i>Portfolio assessment</i>
EVIDENCE	What is assessed?	Writing	Specific, detailed aspects of language for possible future focus in class, e.g., ability to distinguish between fact and opinion in a reading text.	Usually written or spoken language, e.g., peers giving feedback on the extent to which source text content is paraphrased in a student's summary text.	Any aspect of language may be targeted; teacher draws attention to a problem and provides increasingly explicit assistance.	Comprehensive or selective representation of course facets, e.g., samples of writing on different genres throughout a course.
	How is evidence collected?	Teacher corrects or identifies problems with particular grammatical features or a range of features.	A planned test/task designed to elicit predetermined micro-skills.	Peer/s read/listen to each other and offer feedback on aspects of performance (often predetermined and guided by teacher).	Student-teacher or small-group conferences where 'expert' is able to interact with 'novice'; or in computerized environment.	Various samples are collected over a teaching period across a range of planned tasks and at various stages of production.
	When are students assessed?	Between drafts, preferably when there is a possibility of students attending to feedback.	At beginning or end of unit when there is a future possibility of addressing particular weaknesses.	At different stages in teaching cycle when students' developing expertise in a feature/task/genre.	At any stage in teaching cycle.	Collective assessment at end of teaching cycle; individual samples assessed/responded to chronologically.
	Who is assessed?	Individual learner where samples are produced individually	Individual learner or groups	Individual learner	Individual learner or groups	Individual learner where samples are produced individually
	By whom?	Teacher	Teacher	Peer	Teacher/mediator	Teacher

(Continued)

Table 11.2 Continued

DIMENSIONS	Teacher considerations	Written corrective feedback	Diagnostic assessment	Peer assessment	Dynamic assessment	Portfolio assessment
INTERPRETATION	What is the level of attention by teachers and learners?	If implemented with revision processes, can encourage noticing of grammatical gaps. Level of attention may vary.	Sustained and explicit attention to specific aspects of student performance in relation to curriculum.	Sustained if carried out over several iterations of a task or momentary if a short, immediate peer response.	Mediation process sustained according to degree of assistance required.	Sustained across a teaching period.
	Criteria/values/link to standards	Underpinned by expectation that performance is amenable to correction; aspects of language targeted may be drawn from standards.	Detailed criteria drawing out subconstructs of aspect assessed; based on particular models of language skills and their component skills.	Detailed criteria, possibly linked to standards; ability to work with peers may be valued in target domain.	Series of prompts from implicit to explicit mediation; focus of assessment is related to language use prioritized in standards.	Criteria are explicit and detailed and may be related to standards.
USE	How is evidence used?	<ul style="list-style-type: none"> Aggregated information may be used for future teaching To encourage grammatical development through noticing gaps 	<ul style="list-style-type: none"> To adapt teaching/learning activities to cater for areas of weakness To make learners aware of areas of weaknesses and encourage learning 	<ul style="list-style-type: none"> To encourage learning on a number of tasks and language features Learning (both feedback giver and receiver) 	<ul style="list-style-type: none"> Teacher can use individual or group results to inform teaching (and for grading) Socialization into assessment culture 	<ul style="list-style-type: none"> To allocate final grades that are representative of body of work To encourage self-reflection on development
	By whom?	Learner (if CF is attended to) Teacher (e.g., in teaching plan)	Learner (e.g., in goals) Teacher (e.g., in teaching plan) Institution (e.g., in curriculum plan)	Learner (if peer's assessment is attended to and valued) Peer learner (heighten awareness of own language use)	Learner Teacher (e.g., in teaching plan)	Learner (i.e., self-reflection by portfolio author) Teacher (e.g., in making summative judgement)

learner. Written corrective feedback, for example, given its explicit nature, coupled with a revision process would result in noticing if the learner's level of attention (Schmidt, 1990, 2001) is ready for further processing and uptake. Delayed feedback on a portfolio assessment may be less explicit in pointing to particular language problems, in particular if the grading is done at the end of semester. Regardless of the assessment type, it is conceivable that a teacher gives sustained attention to a learner error but the learner barely perceives it and therefore opportunities for noticing and uptake are lost. Similarly, it is possible that a teacher hastily provides a reformulated structure for a particular context, which the learner then uses repeatedly in future as a key feature of discourse structure (Macqueen, 2012). A disparity in attention level is particularly likely when one party prioritizes an aspect of language that is not seen as part of the relevant external standard and therefore not worth bothering with, or when one party is unaware of what the other values in terms of learning outcomes. The most likely scenario for opportunities for learning and acquisition over the long term is if both teacher and learner are engaged in sustained attention to assessment events that are aligned with both external expectations and individual learner needs and characteristics. Continued dialogue between teachers and learners can therefore provide optimal conditions for noticing and uptake.

As can be seen in Table 11.2, an important feature of all the types of assessment represented here is feedback or response. In all these cases, the feedback is intended to be (1) noticed and (2) acted upon. The type of feedback differs, however. Written corrective feedback is highly specific and, to be most effective, requires some pattern detection on the part of the teacher or learner to determine if a grammatical feature is regularly problematic for either an individual or a whole class. If the feedback is focussed on a particular grammatical form, the form is predetermined and should lead to classroom activity that allows feedback action on that form. Peer feedback is constrained in terms of the developing expertise of the individuals involved and their willingness to trust one another's evaluations enough to use it as a basis for revision (Hansen Edwards, 2014). Diagnostic feedback explicitly targets particular, predetermined micro-skills, which are then addressed systematically in future instruction (Alderson, 2005). Portfolio assessment provides a progressive view of development, particularly if it encompasses repeated attempts at the same task type or genre, which enables feedback with a more bird's-eye perspective. In a sense, then, selecting the assessment method is deciding which type of feedback is most effective at different points in a course and most efficiently enhances attention and noticing by students. It is conceivable that in some circumstances, peer-feedback will not be constructively viewed or trusted by students, or that follow-up action on error identification (indirect feedback) is very unlikely (Hansen Edwards, 2014).

To return to the notion of 'assessment cycle' (Rea-Dickins, 2001), it can be seen that most methods are best placed in the learning processes leading to final scoring and reporting. Indeed some may co-occur during the assessment implementation stage. Peer assessment, for instance, may be present in portfolio assessment processes. Written corrective feedback may well be a dynamic process when implemented over time, as a kind of teacher-learner dialogue across different drafts (Macqueen, 2012).

All these issues relate to the degree of alignment between external or imposed standards and classroom practices. Cumming (2009) poses the question: "What should the relationships be among formal language tests, curricula for language learning and pedagogical functions of formative assessment?" (p. 90). It is this issue to which we now turn.

Teaching Tips

- Consider how your assessment methods fit with your general pedagogical approach. For example, if your approach emphasizes the use of authentic communication tasks, try to match this in assessment tasks and criteria.
- Be critical of the evidence you use to determine final grades. Consider to what extent it represents the course content.
- Where possible, include students in the language of the assessments you use. You could increase their understanding of the assessment process by asking them to use the marking criteria to self-assess or to assess an anonymous sample.

Future Directions

Assessment in relation to external standards (including external test measures) is a significant fact of life in the vast majority of ISLA contexts, yet it generally dwells either separately to or on the fringes of theorizing and research about particular approaches to SLA pedagogy. For teachers in contexts where there is significant impact in the classroom from external exams (e.g., Hong Kong) or where reporting against particular standards is mandatory, it can be frustrating and even damaging in some ways to implement instructional approaches that are at odds with the external prescriptions because it may be the external expectations that, in the end, have the more powerful effect on students and other stakeholders. We have tried to emphasize in this chapter that, in practice, classroom-internal learning activities are tied to external expectations in various forms. Therefore, a worthwhile future direction for ISLA research would be to reach beyond the classroom to (1) investigate the effects of external assessment requirements on the implementation of instructional approaches and (2) consider what kind of framework or standards a pedagogical approach might engender and what kinds of assessment practices would be implemented most constructively toward the objectives stipulated by the framework. Rather than considering it outside the remit of approaches to instruction, we would advocate that assessment processes are absolutely integral to teaching practice, and therefore a potent influence on the implementation and efficacy of any teaching approach.

Note

1. The terms ‘frameworks’ and ‘standards’ are used interchangeably in this chapter.

References

- Alderson, C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Alderson, C., & Hamp-Lyons, L. (1996). TOEFL Preparation courses: A study of washback. *Language Testing*, 13(3), 280–297.
- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *Modern Language Journal*, 78(4), 465–483.
- Alshahrani, A., & Storch, N. (2014). Investigating teachers’ written corrective feedback practices in a Saudi EFL context: How do they align with their beliefs, institutional guidelines, and students’ preferences? *Australian Review of Applied Linguistics*, 37(2), 101.

- Angelo, D. (2013). Identification and assessment contexts of Aboriginal and Torres Strait Islander learners of Standard Australian English: Challenges for the language testing community. *Papers in Language Testing and Assessment*, 2(2), 67–102.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102–118.
- Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition*. New York: Routledge.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12(3), 409–431.
- Bitchener, J., & Knoch, U. (2009a). The relative effectiveness of different types of direct written corrective feedback. *System*, 37(3), 322–329.
- Bitchener, J., & Knoch, U. (2009b). The value of a focused approach to written corrective feedback. *ELT Journal*, 63(3), 204–211.
- Bitchener, J., & Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31(2), 193–214.
- Bitchener, J., & Knoch, U. (2010b). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19, 207–217.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14(2), 227–258.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Oxford: Oxford University Press.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Brindley, G. (1998). Describing language development? Rating scales and SLA. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 112–140). Cambridge: Cambridge University Press.
- Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics*, 29, 90–100.
- Danielson, C., & Abrutyn, L. (1997). *An introduction to using portfolios in the classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Davies, A., & LeMahieu, P. (2003). Assessment for learning: Reconsidering portfolios and research evidence. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Innovation an change in professional education: Optimising new modes of assessment: In search of qualities and standards* (pp. 141–169). Dordrecht: Kluwer Academic.
- de Guerrero, M., & Villamil, O. S. (1994). Social-cognitive dimensions of interaction in L2 peer revision. *Modern Language Journal*, 78(4), 484–496.
- de Guerrero, M., & Villamil, O. S. (2000). Activating the ZPD: Mutual scaffolding in L2 peer revision. *Modern Language Journal*, 84(1), 51–68.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Malden, MA: Blackwell.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353–371.
- Ferris, D. (2011). *Treatment of error in second language student writing*. Ann Arbor, MI: University of Michigan Press.
- Fulcher, G. (2003). *Testing second language speaking*. London: Pearson Longman.
- Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Genesee, F., & Upshur, J. A. (1996). *Classroom-based evaluation in second language education*. Cambridge: Cambridge University Press.

Ute Knoch and Susy Macqueen

- Guenette, D., & Lyster, R. (2013). Written corrective feedback and its challenges for pre-service ESL teachers. *Canadian Modern Language Review*, 69(2), 129–153.
- Hamp-Lyons, L., & Condon, W. (2000). *Assessing the portfolio: Principles for practice theory and research*. Cresskill, NJ: Hampton Press.
- Hansen Edwards, J. G. (2014). Peer assessment in the classroom. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 730–750). New York: John Wiley & Sons.
- Harding, L., Alderson, J. C., & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, 32(3), 317–336.
- Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395–420.
- Huhta, A. (2008). Diagnostic and formative assessment. In B. Spolsky & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 469–482). New York: Blackwell.
- Jamieson, J. (2011). Assessment of classroom language learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 768–785). Hoboken, NJ: Taylor & Francis.
- Jang, E. E., Dunlop, M., Park, G., & van der Boom, E. H. (2015). How do young students with different profiles of reading skill mastery, perceived ability, and goal orientation respond to holistic diagnostic feedback? *Language Testing*, 32(3), 359–383.
- Jang, E. E., & Wagner, M. (2014). Diagnostic feedback in the classroom. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 693–711). New York: John Wiley & Sons, Inc.
- Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, 23(2), 12–39.
- Knoch, U. (2009). *Diagnostic assessment of writing: The development and validation of a rating scale*. Frankfurt: Peter Lang.
- Knoch, U. (2015). *Re-examining L2 writing development: Implications for research and pedagogy*. Invited talk at Shanghai Jiao Tong University, School of Foreign Languages.
- Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension of at-risk students. *School Psychology International*, 23(1), 112–127.
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment: Bringing the past into the future. *Journal of Applied Linguistics*, 1(1), 49–74.
- Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2015). Sociocultural theory and second language development. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 207–226). New York: Routledge.
- Lee, J., & Musumeci, D. (1988). On hierarchies of reading skills and text types. *Modern Language Journal*, 72(2), 173–187.
- Leung, C. (2005). Classroom teacher assessment of second language development: Construct as practice. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 869–888). Mahwah, NJ: Lawrence Erlbaum.
- Loewen, S. (2011). Focus on form. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 2, pp. 576–589). Hoboken, NJ: Taylor & Francis.
- Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge: Cambridge University Press.
- Lumley, T. (1993). Reading comprehension sub-skills: Teachers' perceptions of content in an EAP test. *Melbourne Papers in Language Testing*, 2(1), 24–57.
- Macqueen, S. (2012). *The emergence of patterns in second language writing: A sociocognitive exploration of lexical trails* (Vol. 137). Bern: Peter Lang.
- Macqueen, S., Harding, L., & Elder, C. (2011). *An investigation of the ESL VELS student data*. Melbourne: Language Testing Research Centre, University of Melbourne. Study commissioned by Department of Education and Early Childhood Development (DEECD), Government of Victoria.
- Maxwell, G. S. (2004). *Progressive assessment for learning and certification: Some lessons from school-based assessment in Queensland*. Paper presented at the Third Conference of the Association of

- Commonwealth Examination and Assessment Boards, Redefining the Roles of Educational Assessment, Nadi, Fiji, March.
- McKay, P. (2000). On ESL standards for school-age learners. *Language Testing*, 17(2), 185–214.
- McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, 18(4), 333–349.
- McNamara, T., & Elder, C. (2010). Beyond scales. In A. J. Liddicoat & A. Scarino (Eds.), *Languages in Australian education: Problems, prospects and future directions* (pp. 193–202). Newcastle upon Tyne: Cambridge Scholars.
- Meisel, J., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition*, 3, 109–135.
- Nelson, G. L., & Murphy, J. M. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *TESOL Quarterly*, 27(2), 113–131.
- Oscarson, M. (2014). Self-assessment in the classroom. In A. Kunnan (Ed.), *The companion to language assessment* (pp. 712–719). New York: John Wiley & Sons, Inc.
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265–289.
- Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, 43(3), 471–491.
- Poehner, M. E., & Infante, P. (2016). Dynamic assessment in the language classroom. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 275–290). Boston, MA: de Gruyter Mouton.
- Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment. *Language Teaching Research*, 17(3), 323–342.
- Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized Dynamic Assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337–357.
- Purpura, J. (2013). Assessment of grammar. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. New York: Blackwell.
- Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research*, 17(1), 67–89.
- Rastgou, A. (2016). *The impact of teacher feedback on improvement in writing performance in a foreign language context: A longitudinal mixed methods study*. Melbourne: University of Melbourne.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18(4), 429–462.
- Rouhi, A., & Azizian, E. (2013). Peer review: Is giving corrective feedback better than receiving it in L2 writing? *Procedia—Social and Behavioural Sciences*, 93, 1349–1354.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283.
- Sheen, Y., Wright, D., & Modawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adults ESL learners. *System*, 37(2), 556–569.
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5–29.
- Storch, N. (2010). Critical feedback on written corrective feedback. *International Journal of English Studies*, 10(2), 29–46.
- Topping, K. J., Smith, E. F., Swanson, I., & Elliott, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education*, 25(2), 149–169.

Ute Knoch and Susy Macqueen

- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327–369.
- Tsui, A., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147–170.
- Van Beuningen, C. G., de Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62(1), 1–41.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179–200.
- Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom. *Assessing Writing*, 15(1), 3–17.

Section III

Language and Instructed Second Language Acquisition



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Grammar Acquisition

Hossein Nassaji

Background

Grammar is central to language and language learning. As Batstone (1994, p. 4) pointed out, “Language without grammar would be chaotic” and “just as it would be impossible to describe language without seeking out this underlying framework, so it would be impossible to learn a language effectively without drawing on grammar in some way.” Yet, nothing in the field of SLA and language pedagogy has been so controversial as the role of grammar teaching and learning.

This chapter examines the issue of L2 grammar acquisition and the role that grammar instruction plays in the development of grammar knowledge. Grammar knowledge is defined as what learners know about language rules and structures, and the acquisition of grammar is the acquisition of those rules and structures and the ability to use them in a communicative context. The chapter begins with an overview of the theoretical issues and controversies in this area. Then, drawing on findings of current theory and research in SLA, it examines the research evidence related to the role of grammar instruction. A number of research areas are reviewed ranging from studies on the impact of grammar instruction on improving L2 learning in general to those about the role of different types of instruction, the conditions under which instruction is effective, factors influencing the success of instruction, and also whether and how explicit instruction contributes to the development of implicit knowledge. The chapter concludes with the implications of the issues discussed in relation to classroom pedagogy and further research.

Current Issues

One major issue in the field of L2 grammar acquisition concerns the role of grammar teaching and its effect on grammar learning. A persistent debate has been whether grammar can be learned through conscious learning of grammatical rules or whether it should be acquired in the context of meaningful language use. This controversy has been motivated in part by a theoretical debate in the field of cognitive psychology over the role of explicit versus implicit learning and whether learning occurs through conscious manipulation of information or

merely through unconscious processes when people are exposed to input (Bialystok, 1994; N. Ellis, 1993, 1994, 2005; Reber, 1967, 1989, 1993).

Implicit learning is often defined as learning without awareness, taking place when learners are exposed to meaning-focused input, while explicit learning is conscious, taking place mainly through explicit instruction (DeKeyser, 2003, 2005; N. Ellis, 1994; R. Ellis et al., 2009). N. Ellis (2007) pointed out that implicit and explicit learning are functions of separate memory systems, which are located in different areas of the brain. Explicit learning is supported by the neural systems located in the prefrontal cortex responsible for attention, control, and consciousness. Implicit learning, however, involves other areas of the perceptual and motor cortex.

Explicit learning is assumed to lead to explicit knowledge, defined as knowledge that is conscious, learnable, verbalizable, and is “typically accessed through controlled processing when learners experience some kind of linguistic difficulty in using the L2” (R. Ellis, 2006, p. 95). Implicit knowledge, on the other hand, involves no conscious awareness, is procedural, cannot be verbalized, and “is available for use in rapid, fluent communication” (R. Ellis, 2006, p. 95). Implicit knowledge is assumed to occur based on extensive meaning-focused exposure to the target language. Paradis (1994) defined implicit knowledge as knowledge “acquired incidentally” and “stored implicitly” (p. 395).

Central to the debate on the role of grammar learning and teaching is the relationship between the two types of knowledge. This relationship lies at the heart of the discussion of not only grammar teaching and learning but also many other related issues in SLA, including the roles of formal and naturalistic language learning, the differences between first and second language acquisition, and also how children learn differently from adult learners. First language (L1) learners are assumed to acquire L1 grammar mainly implicitly through meaningful exposure to naturalistic input. However, it is unclear whether L2 learners acquire L2 grammar in a similar manner or whether they need instruction. In addition, most researchers agree that the goal of L2 instruction should be the development of implicit knowledge. However, there are the questions of what relationship, if any, exists between explicit and implicit knowledge, and whether and to what extent explicit knowledge assists the development of implicit knowledge. Because in the classroom, explicit knowledge is acquired mainly through explicit instruction, questions have also been raised about the value of explicit instruction and its role in the development of implicit knowledge. These issues are often discussed in the context of what is known as the interface debate, which I will briefly review.

Key Concepts

Implicit learning: Learning without awareness, taking place when learners are exposed to meaning-focused input.

Explicit learning: Learning with awareness, taking place mainly through explicit instruction.

The Relationship Between Explicit and Implicit Knowledge

Traditionally, three positions have been discussed in the field of SLA regarding the relationship between explicit and implicit knowledge: a noninterface, a strong interface, and a weak interface position. These positions differ in the value each assigns to explicit knowledge and hence lead to different recommendations about how to teach grammar.

The noninterface position holds that there is no connection between explicit and implicit knowledge and that the two cannot influence each other. This position was strongly represented in the early 1980s by Krashen's comprehensible input hypothesis, which rested on the distinction between *acquisition* and *learning* and the claim that these two involve independent and unrelated knowledge systems, with acquisition involving unconscious knowledge and learning involving conscious knowledge. Krashen (1982, 1985) argued that grammar instruction leads to conscious learned knowledge and this knowledge cannot turn into sub-conscious acquired knowledge. Thus, he argued that grammar teaching has little impact on the acquisition of language. Similar claims have also been made in the context of Universal Grammar (UG), which suggests that language is mainly learned through the interaction of the principles of UG with the input and not through formal instruction. A number of L2 researchers have applied this perspective to L2 acquisition arguing that similar processes underlie first and second language learning and that if L1 learners do not learn through formal instruction, L2 learners do not need formal instruction either (Cook, 1991; Dulay, Burt, & Krashen, 1982; Schwartz, 1993). Thus, the noninterface position supports teaching approaches that are purely meaning-focused with no attempt to draw learners' attention to form, such as the strong version of communicative language teaching and task-based instruction.

The strong interface position posits that conscious knowledge developed through instruction can turn into implicit or unconscious knowledge. Drawing on information processing theories in cognitive psychology, this perspective maintains that language competence is mainly developed through conscious and declarative knowledge, which becomes proceduralized through ample practice. The strong interface position has gained its support from research that has documented the importance of automaticity in skill learning (DeKeyser, 1997, 2005), and also the neurolinguistic studies that have shown that implicit knowledge is basically a result of proceduralization of explicit knowledge (Paradis, 1994, 2004). Thus, pedagogically, the strong interface position supports approaches that emphasize explicit grammar instruction and practice (an example of which would be the presentation-practice-production, or PPP, model of grammar instruction).

The weak interface position argues that conscious knowledge of grammar can facilitate implicit knowledge but that it does so indirectly through other processes involved in language acquisition. In keeping with the Noticing Hypothesis in SLA (Schmidt, 1993, 1995, 2001), this position holds that explicit knowledge helps learners notice certain language features, which they can subsequently incorporate into their interlanguage if they are developmentally ready (R. Ellis, 1993). Pedagogically, the weak interface position advocates various forms of consciousness raising activities that provide learners with opportunities for attention to form in meaning-focused contexts.

Key Concepts

Noninterface position: There is no connection between explicit and implicit knowledge; explicit knowledge cannot turn into implicit knowledge.

Strong interface position: Explicit knowledge resulting from instruction can become implicit knowledge through ample practice.

Weak interface position: Explicit knowledge facilitates the development of implicit knowledge through promoting other processes (e.g., noticing) that aid acquisition.

Empirical Evidence

Despite various theoretical positions on what kind of knowledge is useful for language acquisition, this is a matter that requires empirical research. A central issue here is the role played by grammar instruction and how the explicit knowledge developed as a result of explicit instruction assists implicit knowledge. Grammar instruction refers to interventional efforts to direct learners' attention to particular grammatical forms. In order to assess the role of instruction, research has focused on a number of key questions. Early research examined whether instruction makes any contribution to language learning in general (e.g., Long, 1983). Subsequent research went beyond this general question, focusing on more specific questions such as what type of instruction is more effective, when it is effective, what factors affect its effectiveness, and what role instruction plays in the development of both explicit and implicit knowledge. In the following sections, I will briefly review the current research evidence in these areas.

The Effectiveness of Grammar Instruction in General

The first question that anyone may ask about grammar instruction is whether or not it has any beneficial effects on L2 acquisition in general. Thus, much of the early research concentrated on this issue. This question was theoretically motivated in part by the position discussed earlier and the claim that explicit and implicit knowledge involves completely distinct mechanisms and that formal instruction does not help the acquisition of language knowledge (Krashen, 1985; Schwartz, 1993).

One of the first studies that formally addressed this question was Long (1983), which examined 12 studies that had compared instructional learning with exposure learning. Long concluded that overall instruction had positive effects on L2 learning as compared to no instruction and that this was true for both children and adults as well as beginner, intermediate, and advanced level learners. R. Ellis (1990, 1994) and Larsen-Freeman and Long (1991) reviewed a number of additional studies and concluded that while instructed language learning did not have major effects on sequences of acquisition, it had facilitative effects on both the rate and the ultimate level of acquisition.

More recent reviews have all arrived at similar conclusions confirming the positive effects of instruction (Doughty, 2001, 2003; Doughty & Williams, 1998; R. Ellis, 2001a, 2001b; Fotos & Hinkel, 2007; Lightbown, 2000; Loewen, 2015; Nassaji & Fotos, 2010; Nassaji & Simard, 2010a, 2010b; Norris & Ortega, 2000, 2001; Russell & Spada, 2006; Spada, 1997; Spada & Tomita, 2010; Williams, 2005). Spada (1997), for example, reviewed a number of L2 classroom and laboratory studies and concluded that form-focused instruction is helpful particularly when incorporated into a communicative context. Norris and Ortega (2000) provided a meta-analysis of 49 form-focused instruction studies and concluded that in general, form-focused instruction produced substantial gain of the target structure knowledge and that the effects were sustained overtime. Spada and Tomita's (2010) meta-analysis of 41 instructional studies also concluded that explicit instruction has a positive effect on L2 acquisition and that this effect is irrespective of the nature of the target structure (see Nassaji, 2016 for a timeline of studies of form-focused instruction).

Thus, there is currently strong empirical evidence for the positive effects of grammar instruction in general. However, even if overall instruction seems to be effective, a closer look at the studies reveals a great deal of variation in results, which then raises the question

of what makes the results variable or what makes instruction sometimes effective and sometimes less effective or ineffective.

Different Types of Instruction

Grammar instruction can encompass a wide range of instructional strategies that differ from one another in important ways such as the degree of explicitness, the mode of instruction, its timing, the degree of planning, obtrusiveness (i.e., interrupting communicative meaning), and the degree to which instruction focuses on input, output, or interaction (Nassaji, 2016). It is quite clear that not all types of instruction are equally effective, but it is unclear what type of instruction is most effective, particularly for the development of implicit knowledge. It is beyond the scope of this chapter to discuss the results of studies in all the related domains. Thus, I will limit my review to a sample of studies in the following areas: explicit versus implicit instruction, focus on form versus focus on forms instruction, input versus output instruction, and the effects of different types of instruction on different types of knowledge.

Explicit Versus Implicit Instruction

One way of classifying grammar instruction is by using the general dichotomy of explicit and implicit instruction. Explicit instruction presents learners with clear information about certain grammatical rules and how they work whereas implicit instruction does not attempt to make learners aware of what they are supposed to learn (R. Ellis, 2008; Hulstijn, 2007; Norris & Ortega, 2000). Research that has compared explicit and implicit instruction, including various forms of explicit and implicit feedback, has generally shown an advantage for explicit instruction over implicit instruction.

For example, in their meta-analysis, Norris and Ortega (2000) compared studies that had used explicit and implicit instruction and concluded that explicit instruction was more effective than implicit instruction. They classified studies as explicit if the treatments involved rule explanation or direct attention to linguistic forms. In the absence of such strategies, the treatment was considered to be implicit. On average, explicit treatments had a considerably larger effect size ($d = 1.13$) than implicit treatments ($d = 0.54$). Spada and Tomita's (2010) meta-analysis also compared explicit and implicit instruction, using the same criteria as Norris and Ortega's to code implicit and explicit instruction. Their results also showed larger effect sizes for explicit instruction than implicit instruction across different measures and target structures.

However, an issue identified by some researchers (e.g., Doughty, 2001, 2003; R. Ellis, 2008) regarding these findings is the excessive use of explicit knowledge tests (i.e., tests of declarative knowledge) as the main measure of language acquisition. Indeed, many of the studies comparing explicit and implicit instruction in Norris and Ortega's meta-analysis had mainly used tests of explicit knowledge rather than those of spontaneous language use. As the researchers pointed out, the majority of studies (about 90%) had used noncommunicative discrete point or metalinguistic tests to measure the role of instruction and only 10% had used measures involving communicative use of language. In addition, about 70% of the studies involved explicit instructional strategies and only 30% involved implicit ones. Furthermore, most studies had operationalized implicit instruction very narrowly as only one type of instruction, whereas explicit instruction often involved a variety of instructional strategies ranging from explanation of rules and practice of those rules to error correction and various

forms of negative feedback. Therefore, their results could be biased toward favoring explicit instruction (Doughty, 2001, 2003; R. Ellis, 2008).

The situation, of course, has changed since Norris and Ortega's study as researchers have begun to develop and use more and more measures that are supposed to tap into implicit knowledge (see the next section). For example, in Spada and Tomita's (2010) meta-analysis, 50% of the studies had used measures of implicit knowledge, including tests of free production and spontaneous language use. The fact that their findings still showed a greater effect for explicit instruction suggests that explicit instruction can be overall more effective than implicit instruction.

Key Concepts

Explicit instruction: Presents learners with clear information about target grammatical rules.

Implicit instruction: Does not provide learners with explicit information about the target rules.

Focus on Form Versus Focus on Forms

Another widely cited distinction that has had a considerable impact on our understanding of grammar instruction is the one that Long (1991) drew between *focus on form* and *focus on forms*. Focus on forms is the traditional structure-based instruction in which language is segmented into discrete items and then presented to learners in an isolated and de-contextualized manner. Focus on form, on the other hand, involves drawing learners' attention to linguistic forms "as they arise incidentally in lessons whose overriding focus is on meaning or communication" (Long, 1991, pp. 45–46). An example of focus on form instruction is interactional feedback, such as recasts, which provide learners with the correct reformulation of their errors in the course of meaning-focused interaction. Some examples of focus on forms instruction are isolated grammar exercises such as pattern drills, fill-in-the-blanks, or other activities typical of the traditional grammar translation method.

As far as research is concerned, there is evidence that instruction that occurs in a meaning-focused context is more effective than instruction that focuses on grammatical forms in isolation (Doughty, 2003; R. Ellis, 2008; Lightbown & Spada, 1993; Nassaji & Fotos, 2004; 2010; Spada, 1997). Reviewing a number of classroom studies, Lightbown and Spada (1993) concluded:

form-focused instruction and corrective feedback provided within the context of a communicative program are more effective in promoting L2 learning than programs which are limited to an exclusive emphasis on accuracy on the one hand or an exclusive emphasis on fluency on the other.

p. 105

However, despite the overall positive effect of an integration of attention to form into meaning-focused classrooms, studies that have more directly compared focus on form with focus on forms instruction have not found a clear difference between the two. For example, as part of their meta-analysis, Norris and Ortega (2000) compared focus on form studies with focus on forms studies (that is, those that taught linguistic forms in a meaning-focused context

versus those that focused on forms outside of communicative situations) and found both to be equally effective, yielding similar effect sizes (focus on form, $d = 1.92$; focus on forms, $d = 1.47$). Of course, although Norris and Ortega distinguished between focus on form and focus on forms studies, most studies classified as focus on form involved some kind of explicit instruction, which makes their conclusion difficult to interpret.

However, a few more recent studies have also compared the two types of instruction (de la Fuente, 2006; Shintani, 2013, 2015; Shintani & R. Ellis, 2010; Valeo, 2013) and have also not found a clear difference between the two types of instruction. Shintani (2013), for example, examined the difference between focus on form and focus on forms and found that both treatment groups showed improvement in learning English nouns but the focus on form group was better at acquiring adjectives (see also Shintani & R. Ellis, 2010). Valeo (2013) compared focus on form with meaning-focused instruction on learning two grammatical targets: the present conditional and the simple past tense of English. Pretest–posttest measures showed significant gains for both types of instruction. De la Fuente (2006) compared focus on forms (operationalized as the PPP method of teaching) with focus on form in task-based instruction among adult university students and found that task-based instruction was more effective than PPP lessons. However, in this study, it was the explicitness of focus on form that made it more effective than task-based lessons. These findings suggest that both focus on form and focus on forms instruction can be effective depending on how each is provided. For example, focus on forms can be an effective approach if learners also practice language forms in communicative tasks (R. Ellis, 2006).

Having said that, there has been a great variation in studies as to what constitutes a focus on form or a focus on forms instruction. Therefore, we cannot draw a generalizable conclusion based on the results of these studies. For example, Shintani's (2013) study of a focus on form versus a focus on forms instruction was actually a comparison of comprehension and production-based lessons. Furthermore the production-based lessons (which were considered as focus on forms lessons) included recasts, a reactive type of focus on form. Therefore, those lessons could not necessarily be considered as focus on forms instruction alone. In order to be able to compare reliably the two types of instruction, studies need to isolate instances of focus on form from focus on forms, and in doing so, there is a need for standard and explicit criteria to distinguish between the two types of instruction. However, this has not yet been the case (R. Ellis, 2016).

Teaching Tip

Make sure to include some kind of attention-to-form or consciousness-raising activities into the design of communicative lessons. This can be done, for example, by explaining certain grammatical forms before or after a communicative activity, by using feedback during interaction, or by using input enhancement strategies that highlight grammatical forms in the course of meaning-focused discourse.

Input-Based and Output-Based Instruction

Grammar instruction can also be categorized in terms of whether the focus is on input or output. Input-based instruction refers to instructional strategies that involve the use or the processing of input. This approach is based on the assumption that learners' attention can

be drawn to grammatical forms through activities whose aim is to understand input for meaning. Output-based instruction refers to instruction that draws attention to grammatical structures through eliciting and practicing learners' output.

One type of input-based instruction that has received much attention is Processing Instruction (PI). PI is a particular approach to grammar instruction that is based on how learners process the input data (VanPatten, 2002, 2004). This perspective holds that instruction is beneficial if learners are helped to attend to linguistic forms when learners are processing input for meaning (VanPatten, 2015). A number of studies have compared the effectiveness of PI with those of both meaning-focused instruction and the traditional output-based instruction (e.g., Allen, 2000; Benati, 2001, 2004, 2005; Benati & Lee, 2008; Cadierno, 1995; Cheng, 2002; J. Lee & Benati, 2013; Morgan-Short & Bowden, 2006; VanPatten & Cadierno, 1993; Wong, 2004). Overall, their results have shown supportive evidence for PI. However, they have also shown that the effectiveness of PI depends on a number of factors, including the kind and complexity of the target structure and the type of skill measured. In a recent review of the studies of PI, DeKeyser and Botana (2015) concluded that PI might be more effective for promoting comprehension skills whereas production-based instruction might be more effective for promoting production skills (see also DeKeyser & Sokalski, 2001). They also noted that the results of studies comparing PI with output-based instruction depended on how input or output-based instruction was operationalized. A few other points about PI studies are that they often have used input activities in the form of decontextualized sentences in combination with explicit instruction, and have tested their effects via measures favoring explicit rather than implicit knowledge. Therefore the extent to which PI can help learners to use language spontaneously in communicative contexts is unclear.

Examples of input-based instruction also include various forms of input enhancement techniques such as textual enhancement and input flood. The aims of these strategies are to raise learners' attention to form by rendering input perceptually more salient. Textual enhancement aims to achieve this by highlighting certain aspects of the input by means of various typographic devices, such as bolding, underlining, and italicizing in written input, or various acoustic devices such as added stress or repetition in oral input (Nassaji & Fotos, 2010). Input flood involves the provision of numerous examples of a certain target form in the input (either oral or written). The assumption here is that frequent instances of the same target form make the form perceptually salient, drawing the learners' attention to that form (Nassaji & Fotos, 2010).

Studies that have examined the effectiveness of textual enhancement and input flood (e.g., Alanen, 1995; Han, Park, & Combs, 2008; Hernandez, 2011; Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995; S. K. Lee & Huang, 2008; Simard, 2002, 2009; Trahey & White, 1993; J. White, 1998; L. White, 1991) have produced inconclusive results. Some studies, for example, have shown an overall positive effect (Jourdenais et al., 1995; S. K. Lee, 2007; Simard, 2009; J. White, 1998), while others have reported limited effects for these strategies (e.g., Alanen, 1995; Leow, 1997; Overstreet, 1998; Wong, 2003).

Alanen (1995), for example, examined the effects of textual enhancement versus explicit instruction on the acquisition of Finnish locative features and consonant gradation, and found that the textual enhancement group benefited most from the treatment. However, the group who received explicit instruction outperformed the group who did not receive such instruction. White (1998) examined the effects of textual enhancement and input flood on learning third person singular possessives in English among French-speaking children and found some effects on learners' noticing of the targeted form but not on improving learning. Leow (2001) found no advantage for enhanced text over unenhanced text for learning

Spanish formal imperatives. Finally, S.K. Lee and Huang (2008) provided a meta-analysis of 16 primary studies of textual enhancement and found small effects of input enhancement strategies on L2 learning.

Of course, some studies have shown greater effects for the more explicit forms of textual enhancement and input flood. Simard (2009) found that textual enhancement was most effective on noticing the target structure (English plural markers) when their salience was enhanced through a combination of formats. Williams and Evans (1998) found that input flood plus explicit instruction was more effective than implicit input flood on learning participial adjectives (see also Hernandez, 2011; L. White, 1991, for the positive effect of input flood plus explicit instruction). These findings confirm the role of explicitness in instruction and its positive effects on L2 learning.

Taken together, studies examining the effectiveness of textual enhancement including input flood have shown varying results. While most of the studies suggest an overall positive effect for such techniques on noticing, they do not provide proof of learning. Such findings are not surprising because textual enhancement simply provides learners with the correct models of the language or what is known as positive evidence. It does not provide learners with information about what is incorrect in a given language, or what has been called negative evidence. Thus, although such strategies may enhance the salience of the target structure and hence may result in noticing the form, textual enhancement may not lead to a deeper level of cognitive processing needed for acquisition. This could also be because they simply involve comprehension and not production.

The Effects of Different Types of Instruction on Different Types of Knowledge

Another important question regarding the role of grammar instruction concerns what type of knowledge benefits most from what type of instruction. If we agree that what underlies spontaneous use of language in communicative contexts is primarily implicit knowledge, then the question arises as to what role explicit instruction plays in the development of implicit knowledge. A few recent studies have examined the differential effects of implicit and explicit instruction on explicit and implicit knowledge.

One of the studies that examined the effects of explicit instruction on implicit knowledge is that of R. Ellis (2002), which analyzed 11 studies of grammar instruction that had used communicative free production as a measure of implicit knowledge. R. Ellis concluded that explicit instruction contributed to the acquisition of implicit knowledge. The analysis also highlighted the importance of two factors that mediated the success of instruction: the kind of target structure and the extent of instruction. That is, instruction was more likely to be effective for simple target features (e.g., verb forms, articles) than more complex ones (e.g., English passive forms) particularly when the instruction consisted of several hours of instruction spread over several weeks rather than 1 or 2 hours of instruction. R. Ellis, Loewen, and Erlam (2006) compared the usefulness of explicit metalinguistic explanation and implicit recasts on the development of English past tense *-ed* among low-intermediate L2 learners. Using tests of both implicit (an oral imitation test) and explicit knowledge (an untimed grammaticality judgment test and a metalinguistic knowledge test), they found that explicit metalinguistic feedback contributed to the development of both implicit and explicit knowledge.

Andringa, de Glopper, and Hacquebord (2011) conducted a classroom study with English learners of Dutch as an L2. Learners received either explicit or implicit instruction on two

types of Dutch structures: the degrees of comparison and verb-final position in subordinate clauses. An untimed grammaticality judgment task and a free written response task were used to measure explicit and implicit knowledge respectively. The study found that both explicit and implicit instruction facilitated the use of the target structures in free response written tasks, indicating that the two types of instruction promoted implicit knowledge. However, for one of the target structures (the degrees of comparison), explicit instruction was more effective than implicit instruction when there was a similarity between learners' L1 and the L2, suggesting a mediating role of the learner's L1.

In their meta-analysis, Spada and Tomita (2010) also found that explicit instruction contributed to the development of both explicit knowledge (as measured by controlled tasks such as metalinguistic judgment and multiple choice tests) as well as implicit knowledge (as measured by free-response measures such as picture description and information gap tasks). Indeed, they found that the effect of explicit instruction on implicit knowledge yielded the largest effect size for complex target structures. Of course, most studies in their meta-analysis had compared the effect of instruction with a control condition and few had included a comparison of explicit and implicit instruction (Andringa & Curcic, 2015). Therefore, the results could have been affected by the way the studies were analyzed in the meta-analysis.

Teaching Tips

- Provide opportunities for the development of both explicit and implicit knowledge, but do not assume that explicit knowledge will be converted automatically into implicit knowledge.
- Be aware that learning a language is a gradual process that takes time. Although instruction is important for raising learners' attention to form, the key to the development of implicit knowledge is continual exposure to meaningful input and practice. Therefore, provide opportunities for repeated use of the target grammatical forms in meaningful communicative contexts.

Pedagogical Implications

In this section, I will discuss some of the pedagogical implications that can be drawn from the issues examined while considering the factors and the conditions that can affect the role of instruction. Looking at the various issues discussed, one point that stands out is that explicit grammar instruction is beneficial for the development of L2 knowledge, including implicit knowledge. As noted earlier, Spada and Tomita's (2010) meta-analysis indicated that explicit instruction was more effective than implicit instruction irrespective of the type of target structure, and Norris and Ortega (2001, p. 203) concluded that "empirical findings indicate that explicit instruction is more effective than implicit instruction." Explicit instruction leads to explicit knowledge, but it also facilitates the development of implicit knowledge by making learners more likely to notice the forms in subsequent input and gradually internalize them. This then can suggest that teachers should include some form of explicit instruction into the design of their communicative lessons when needed.

However, although explicit instruction has been shown to be effective, the relationship between instruction and learning is complex and the benefits of instruction may not occur unless it takes place under suitable conditions. Explicit instruction can also take various forms. It can occur in the form of traditional grammar-based lessons or it can be integrated

into a communicative context. As noted earlier, the results of form-focused studies indicate that instruction is overall most effective when it is incorporated into a meaning-focused context, which suggests that to maximize learning, teachers should attempt to combine a focus on grammar with a focus on meaning. Such a combination can be achieved in different ways such as by using various kinds of communicative and grammar consciousness-raising tasks, using activities that provide opportunities for both guided and free practice, and feedback on learner errors in the course of a communicative activity (see Nassaji, 2015; Nassaji & Fotos, 2010 for a detailed discussion of these activities and their classroom application).

In classroom contexts, opportunities for focus on language forms can also be created through what have been called problem-solving *grammar tasks* (Nassaji, 1999; Nassaji & Fotos, 2010). In such tasks, learners are presented with language activities that illustrate some language structures. Learners are then asked to work in pairs or small groups and reflect on language form and try to discover the grammatical rules underlying the structure (e.g., Fotos, 1993, 1994; Fotos & R. Ellis, 1991). Such tasks may be a more effective option than traditional grammar instruction as they may help learners better understand form–meaning relationships (Nassaji & Fotos, 2010). Because the tasks are completed collaboratively and learners also use the target language to communicate about language, such tasks may provide an effective way of integrating a grammar focus into communicative tasks.

As mentioned before, instruction may not assist language acquisition all the time and it should meet certain criteria to be effective. In this respect, one of the major factors in determining the success of instruction is the learners' developmental readiness. R. Ellis (1993, 1994) pointed out that explicit instruction is helpful only if learners are developmentally ready to acquire the target structure. He maintained:

[T]he learner's existing knowledge constitutes a kind of filter that sifts explicit knowledge and lets through only that which the learner is ready to incorporate into the interlanguage system. In other cases, however—when the focus of the instruction is a grammatical property that is not subject to developmental constraints—the filter does not operate, permitting the learner to integrate the feature directly into implicit knowledge.

R. Ellis, 1994, pp. 88–89

This suggests that teachers should consider learners' developmental readiness and try to teach the forms that accord with learners' developmental levels. However, one issue with the idea of developmental readiness is the difficulty of tailoring instruction to each individual learner's developmental level, particularly if the classroom consists of mixed-ability learners. The teacher may also not know exactly the developmental level of each learner in relation to specific target structures. Having said that, teachers can still find ways to make instruction suitable to students' levels. One way would be by using a wide range of activities or instructional strategies to present grammatical forms. If the teacher uses a variety of grammar techniques, students may have opportunities to benefit from the instruction based on their own individual needs and interests. Another possibility is to make use of interactional feedback in the course of communicative activities. Interactional feedback provides an alternative approach to error correction typical of traditional grammar-based approaches, in which correction is often given in a decontextualized manner (see Nassaji, 2015; and also Nassaji, 2016). Because interactional feedback takes place when learners make an error while communicating, attention to form takes place at the time when learners need it. Also, because students are engaged in communicative interaction, the feedback helps learners attend to form at the time when they are

processing form for meaning (e.g., R. Ellis & Sheen, 2006; Gass, 1997; Long, 1996; Mackey, 1999; Nassaji, 2016).

In addition to learners' developmental levels, there are other factors that may affect the effect of instruction, which the teacher should consider. One factor, for example, is the nature of the target structure. It is important to note that not all grammatical forms are the same or learned in the same manner. For example, grammatical forms may differ in their linguistic complexity, transparency of form-function mapping, salience, and frequency, and these differences may influence the effects of instruction. Some grammatical forms, for example, are nonsalient (such as certain function words) and therefore may be harder to notice in the input. Learners may also have difficulty in establishing form-function mapping for certain grammatical features such as particles or inflections (N. Ellis & Collins, 2009). In such cases, instruction may become more beneficial as it helps learners to attend to these forms.

The frequency of the target form is yet another related factor. Grammatical forms that have low frequency in the input may be harder to notice due to their rareness (Lightbown, 1992). Instruction that draws learners' attention to these target forms can be beneficial as it may help learners become aware of such forms, which may not otherwise be noticed.

Instruction may also be beneficial or even required for grammatical features with little communicative value (i.e., forms that do not contribute much to the overall meaning) (VanPatten, 2004). If the forms are of little communicative value, learners may not pay attention to them when processing input for meaning. They may either focus on meaning without paying adequate attention to form, or they may pay attention to form without adequately processing meaning. Part of the reason for this may be that L2 learners have limited attentional capacity and therefore may have difficulty attending to both form and meaning at the same time (VanPatten, 2002).

Last but not least, the effect of instruction may also be mediated by various individual learner differences such as learners' age, aptitude, personality characteristics, language proficiency, motivation, attitudes toward learning, cultural backgrounds, and L1. Thus, teachers should be aware of these factors and attempt to take them into account as much as possible. For instance, a deductive approach to teaching grammar may work better for students who are accustomed to a deductive approach or those who have a deductive learning style than those who do not.

Teaching Tips

- Students learn a target form when they are linguistically and cognitively prepared to acquire it. Thus, teachers should target features that learners are ready to learn.
- Consider the nature of the target structure as an important factor in designing your grammar lessons.
- Learners are different and learn differently. Thus, take into account the various individual learner differences that can mediate the effect of instruction.

Further Directions

As can be noted, there is considerable theoretical and empirical research in SLA on the role of grammar learning and instruction. The findings of this research have contributed much

insight into the processes involved and have also led to new ideas and different ways to conceptualize and understand the complexities involved in grammar learning and teaching. Despite considerable insight from this research, however, there are still many questions that have remained unanswered or partially answered.

As reviewed, a number of studies have examined the effectiveness of different types of instruction. For example, there is evidence that explicit instruction can be more effective than implicit instruction. However, we still know little about how the two types of instruction interact and in what ways they contribute to the development of either explicit or implicit knowledge. In addition, much remains to be known about the most valid and reliable ways of measuring different types of knowledge.

For instance, one measure used in most studies to assess explicit knowledge is the untimed grammaticality judgment test. However, the results of such measures should be treated with caution as it is unclear what learners do when they perform a grammaticality judgment task (R. Ellis, 2005). Based on a review of the literature in this area, Hedgcock (1993) argued that it is hard to know what processes are involved in decisions regarding grammaticality or ungrammaticality of an utterance in a grammaticality judgment, what kind of knowledge learners rely on, and what aspect of that knowledge may affect their decision. Many factors may also influence learners' judgments including the complexity of the task, the target structure, and learners' level of language proficiency.

As for implicit knowledge, research has used measures such as free L2 production, oral imitation tasks, or timed grammaticality judgment tests. However, we cannot assume that these measures are the same or would tap equally into the same type of knowledge. In addition, many of the L2 production tasks used to measure implicit knowledge in research are not entirely free and spontaneous. For this purpose, for instance, research has often used picture-cued tasks designed to elicit certain target structures. Thus, they are to some degree planned and controlled, which may then cause learners to process the form more consciously by relying on their explicit knowledge when completing the tasks.

The preceding issues suggest that research should not only continue to examine the role of various explicit and implicit instruction strategies but also try to develop ways in which implicit and explicit knowledge can be more reliably measured.

As noted earlier, research has suggested that grammar instruction is helpful if learners have reached the developmental level required to learn the target structure. However, it is unclear what this developmental threshold is for different target structures. As mentioned, for instruction to be effective, the proper choice of the target structures is critical. Also, we do not yet know how and in what ways developmental readiness interacts with other factors (e.g., learners' L1) that may influence the effectiveness of instruction. Furthermore, although developmental research has been conducted on certain English language structures such as question formation, negation, relative clauses, and certain morphological features, we do not yet know about many other language structures and the extent to which they follow predictable developmental sequences. And for those structures that do follow a fixed route, it is unclear how classroom instruction should be designed to target those structures most effectively.

With respect to the role of the target structure, a number of studies have compared the effect of instruction on complex versus simple structures. However, the results are mixed. While some have found that instruction is more effective for simple structures (e.g., Robinson, 1996), others have found explicit instruction to be equally effective for simple and complex structures (e.g., Housen, Pierrard, & Van Daele, 2005 see also Spada & Tomita, 2010). Williams and Evans (1998) found a greater effect of explicit instruction on less complex

structures but an equal effect for explicit and implicit instruction for the more complex structures. One possible reason for such differences could be the difficulty in defining complexity, and therefore various studies have operationalized this variable differently. This variation then indicates the need for identifying more consistent and relevant criteria as well as research to test those criteria systematically across studies. Of course, we should note that there is no straightforward relationship between linguistic complexity and learning. Some features may be linguistically simple but difficult to explain such as English definite and indefinite articles (*the* and *a*). Others may be both linguistically simple and also easy to explain but difficult to fully acquire, such as the third person singular -s.

Finally, as reviewed, there is evidence that grammar instruction that integrates attention to form in a communicative context is more effective than instruction that targets grammatical structures out of context. However, it is not yet clear how such integration can be achieved to be maximally effective for different target structures and for different learners (from different levels of language proficiency or L1 backgrounds). Furthermore, although some studies have compared focus on form and focus on forms instruction, very few studies have compared the different ways of providing focus on form instruction. Spada, Jessop, Suzuki, Tomita, and Valeo (2014) have recently compared two kinds of communicative attention to form (one that occurred within a communicative activity and attention to form that occurred separate from it) and found no difference between the two. They attributed this lack of difference to the idea that both methods combined focus on form and focus on meaning. However, research has also shown that the timing of focus on form may have differential effects. Kim (2014), for example, found that focus on form prior to meaning-focused instruction was more successful than focus on form that occurred with some delay after the meaning-focused instruction. Different types of focus on form may also be differentially effective for different learners and different types of target structures. These possibilities suggest that more research is needed in these areas.

References

- Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language acquisition* (pp. 259–302). Honolulu: University of Hawai'i.
- Allen, L. Q. (2000). Form–meaning connections and the french causative: An experiment in processing instruction. *Studies in Second Language Acquisition*, 22, 69–84.
- Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing. *Studies in Second Language Acquisition*, 37, 237–268.
- Andringa, S., de Glopper, K., & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning*, 61, 868–903.
- Batstone, R. (1994). *Grammar*. Oxford: Oxford University Press.
- Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research*, 5, 95–127.
- Benati, A. (2004). The effects of processing instruction and its components on the acquisition of gender agreement in Italian. *Language Awareness*, 13, 67–80.
- Benati, A. (2005). The effects of processing instruction, traditional instruction and meaning—output instruction on the acquisition of the English past simple tense. *Language Teaching Research*, 9, 67–93.
- Benati, A., & Lee, J. (2008). *Grammar acquisition and processing instruction: Secondary and cumulative effects*. Clevedon: Multilingual Matters.

- Bialystok, E. (1994). Representation and ways of knowing: Three issues in second language acquisition. In N. Ellis (Ed.), *Explicit and implicit learning of languages* (pp. 549–569). London: Academic Press.
- Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *Modern Language Journal*, 79, 179–193.
- Cheng, A. (2002). The effects of processing instruction on the acquisition of *ser* and *estar*. *Hispania*, 85, 900–909.
- Cook, V. (1991). *Second language learning and language teaching*. London: Edward Arnold.
- DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195–221.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford: Blackwell.
- DeKeyser, R. (2005). Implicit and explicit learning. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 314–348). Mahwah, NJ: Lawrence Erlbaum Associates.
- DeKeyser, R., & Botana, G. (2015). The effectiveness of processing instruction in L2 grammar acquisition: A narrative review. *Applied Linguistics*, 63, 290–305.
- DeKeyser, R., & Sokalski, K. (2001). The differential role of comprehension and production practice. *Language Learning*, 51, 81–112.
- de la Fuente, M. J. (2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research*, 10, 263–295.
- Doughty, C. (2001). Cognitive underpinning of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206–257). Cambridge: Cambridge University Press.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256–310). Oxford: Blackwell.
- Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. New York: Oxford University Press.
- Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5, 289–318.
- Ellis, N. (1994). *Implicit and explicit learning of languages*. London/San Diego: Academic Press.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305–352.
- Ellis, N. (2007). Implicit and explicit knowledge about language. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education, second edition, Vol. 6: Knowledge about language* (pp. 119–132). New York: Springer.
- Ellis, N., & Collins, L. (2009). Input and second language acquisition: The roles of frequency, form, and function introduction to the special issue. *Modern Language Journal*, 93, 329–335.
- Ellis, R. (1990). *Instructed second language acquisition: Learning in the classroom*. Oxford: Blackwell.
- Ellis, R. (1993). The structural syllabus and second language-acquisition. *TESOL Quarterly*, 27, 91–113.
- Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 79–114). San Diego: Academic.
- Ellis, R. (2001a). *Form-focused instruction and second language learning*. Malden, MA: Blackwell.
- Ellis, R. (2001b). Introduction: Investigating form-focused instruction. *Language Learning*, 51, 1–46.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? A review of the research. *Studies in Second Language Acquisition*, 24, 223–236.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172.
- Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40, 83–107.

- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research*, 20, 405–428.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–369.
- Ellis, R., Loewen, S., Erlam, R., Philp, J., Elder, C., & Reinders, H. (Eds.). (2009). *Implicit and explicit knowledge in second language learning and teaching*. Clevedon: Multilingual Matters.
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28, 575–600.
- Fotos, S. (1993). Consciousness-raising and noticing through focus on form—grammar task-performance versus formal instruction. *Applied Linguistics*, 14, 385–407.
- Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28, 323–351.
- Fotos, S., & Ellis, R. (1991). Communicating about grammar: A task-based approach. *TESOL Quarterly*, 25, 605–628.
- Fotos, S., & Hinkel, E. (2007). Form-focused instruction and output for second language writing gains. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education* (pp. 131–143). Oxford: Oxford University Press.
- Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Lawrence Erlbaum.
- Han, Z., Park, E., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics*, 29, 597–618.
- Hedgcock, J. (1993). Well-formed vs. Ill-formed strings in L2 metalingual tasks: Specifying features of grammaticality judgements. *Second Language Research*, 9, 1–21.
- Hernandez, T. (2011). Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research*, 15, 159–182.
- Housen, A., Pierrard, M., & Van Daele, S. (2005). Rule complexity and the efficacy of explicit grammar instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 235–269). Amsterdam: Mouton de Gruyter.
- Hulstijn, J. (2007). Fundamental issues in the study of second language acquisition. *EUROSLA Yearbook*, 7, 191–191.
- Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual enhancement promote noticing? A think-aloud protocol analysis. In R. Schmidt (Ed.), *Attention and awareness in second language learning* (pp. 183–216). Technical Report No. 9. Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Kim, J. (2014). Timing of form-focused instruction and development of implicit vs. explicit knowledge. *English Teaching*, 69, 123–147.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Pergamon Press.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Oxford: Pergamon Press.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. London: Longman.
- Lee, J., & Benati, A. (Eds.). (2013). *Individual differences and processing instruction*. London: Equinox.
- Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning*, 57, 87–118.
- Lee, S. K., & Huang, H. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition*, 30, 307–331.
- Leow, R. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning*, 8, 151–182.
- Leow, R. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania*, 84, 496–509.
- Lightbown, P. (1992). Getting quality input in the second/foreign language classroom. In C. Kramsch & S. McConnell-Ginet (Eds.), *Text and context: Cross-disciplinary and cross-cultural perspectives on language study* (pp. 187–197). Lexington, MA: D. C. Heath.

- Lightbown, P. (2000). Anniversary article: Classroom SLA research and second language teaching. *Applied Linguistics*, 21, 431–462.
- Lightbown, P., & Spada, N. (1993). *How languages are learned*. Oxford/New York/Toronto: Oxford University Press.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Long, M. (1983). Does second language instruction make a difference. *TESOL Quarterly*, 17, 359–382.
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. DeBot, R. Ginsberge, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: John Benjamins.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic Press.
- Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21, 557–587.
- Morgan-Short, K., & Bowden, H.W. (2006). Processing instruction and meaningful output-based instruction: Effects on second language development. *Studies in Second Language Acquisition*, 28, 31–65.
- Nassaji, H. (1999). Towards integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *Canadian Modern Language Review*, 55, 385–402.
- Nassaji, H. (2015). *The interactional feedback dimension in instructed second language learning: Linking theory, research, and practice*. London: Bloomsbury.
- Nassaji, H. (2016a). Research timeline: Form-focused instruction and second language acquisition. *Language Teaching Research*, 49, 35–62.
- Nassaji, H. (2016b). Anniversary article: Interactional feedback in second language teaching and learning: A synthesis and analysis of current research. *Language Teaching Research*, 20, 535–562.
- Nassaji, H., & Fotos, S. (2004). Current developments in research on the teaching of grammar. *Annual Review of Applied Linguistics*, 24, 126–145.
- Nassaji, H., & Fotos, S. (2010). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. London: Routledge.
- Nassaji, H., & Simard, D. (Eds.). (2010a). Current developments in form-focused interaction and L2 acquisition. [Special Issue]. *Canadian Modern Language Review*, 66(6).
- Nassaji, H., & Simard, D. (2010b). Introduction: Current developments in form-focused interaction and L2 acquisition. *Canadian Modern Language Review*, 66, 773–778.
- Norris, J.M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Norris, J.M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51, 157–213.
- Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics*, 2, 229–258.
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393–419). San Diego: Academic Press.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (Vol. 18). Amsterdam, The Netherlands: John Benjamins.
- Reber, A. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.
- Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology-General*, 118, 219–235.
- Reber, A. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. Oxford: Clarendon Press.

- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 131–164). Amsterdam: John Benjamins.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206–226.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: University of Hawai'i Press.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.
- Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15, 147–163.
- Shintani, N. (2013). The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of L2 vocabulary by young beginning-level learners. *TESOL Quarterly*, 47, 36–62.
- Shintani, N. (2015). The incidental grammar acquisition in focus on form and focus on forms instruction for young beginner learners. *TESOL Quarterly*, 49, 115–140.
- Shintani, N., & Ellis, R. (2010). The incidental acquisition of English pluralŋs by Japanese children in comprehension-based and production-based lessons. *Studies in Second Language Acquisition*, 32, 607–637.
- Simard, D. (2002). Textual enhancement—where do we start and where are we headed? *Canadian Modern Language Review*, 59, 236–263.
- Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System*, 37, 124–135.
- Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, 29, 1–15.
- Spada, N., Jessop, L., Suzuki, W., Tomita, Y., & Valeo, A. (2014). Isolated and integrated form-focused instruction: Effects on different types of L2 knowledge. *Language Teaching Research*. Advance online publication. doi:10.1177/1362168813519883
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta analysis. *Language Learning*, 60, 263–308.
- Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition*, 15, 181–204.
- Valeo, A. (2013). The integration of language and content: Form-focused instruction in a content-based language program. *Canadian Journal of Applied Linguistics*, 16, 25–50.
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52, 755–803.
- VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Lawrence Erlbaum.
- VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics in Language Teaching*, 53, 91–109.
- VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15, 225–244.
- White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85–113). Cambridge: Cambridge University Press.
- White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7, 133–161.
- Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 673–691). Mahwah, NJ: Lawrence Erlbaum Associates.

- Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139–155). Cambridge: Cambridge University Press.
- Wong, W. (2003). Textual enhancement and simplified input: Effects on L2 comprehension and acquisition of non-meaningful grammatical form. *Applied Language Learning*, 13, 109–132.
- Wong, W. (2004). The nature of processing instruction. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 33–63). Mahwah, NJ: Lawrence Erlbaum.

Acquisition of L2 Pragmatics

Kathleen Bardovi-Harlig

Background

This chapter investigates the intersection of pragmatics, “the study of language from the point of view of users, especially of the choices they make, the constraints they encounter in using language in social interaction and the effects their use of language has on other participants in the act of communication” (Crystal, 1997, p. 301); and instructed SLA (ISLA), a theoretically and empirically based field of academic inquiry that aims to understand how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one’s own (Loewen & Sato, this volume). Said less formally, the study of pragmatics in ISLA is the study of instructional and learner variables involved in facilitating the classroom learning of how to say what to whom when, in a second or foreign language (Bardovi-Harlig, 2013).

The study of pragmatics is traditionally held to encompass at least five main areas: deixis, conversational implicature, presupposition, speech acts, and conversational structure (Levinson, 1983). In second language (L2) research, pragmatics has also included the choice of address terms, conversation management (including turn-taking), and the use of pragmatic routines and conventional expressions, and has not yet turned its attention to deixis or presupposition (Bardovi-Harlig, 2010). Research in pragmatics often distinguishes between pragmalinguistics—the language resources speakers use for pragmatic purposes—and sociopragmatics—the rules that guide use of language in society and in context.

Key Concepts

Pragmalinguistic Knowledge: Knowledge of the language resources speakers use for pragmatic purposes, for example, knowing that ability questions in English can be used for requests such as “Can you pass the salt?” when a speaker would like the salt.

Sociopragmatic Knowledge: Knowledge of the rules that guide use of language in society and in context. Sociopragmatic knowledge includes knowing what speech acts are appropriate in what contexts, such as thanking the instructor at the end of a meeting during office hours (rather than offering an apology for taking the instructor's time).

The primary instructional target of ISLA studies mirrors the most common area of research in pragmatics at large, namely speech acts. Instruction has targeted a range of speech acts including apologies, refusals, compliments and compliment responses, requests, complaints, suggestions, agreements, disagreements, and thanking. Targets other than speech acts are less well represented in ISLA, but nevertheless include a variety of constructs:

- Lexical modifiers, downgraders, and speech act modifiers (Barekat, 2013; Nguyen, 2013; Safont Jordá, 2003; Safont Jordá & Alcón Soler, 2012);
- Back channel signals and reactive expressions (Sardegna & Molle, 2010);
- Pragmatic routines (Bardovi-Harlig, Mossman, & Vellenga, 2015b);
- Speech events including handling customer complaints (Trosborg & Shaw, 2008), job interviews (Louw, Derwing, & Abbott, 2010), and argumentation (Németh & Kormos, 2001);
- Address terms in German (Kinging & Belz, 2005) and French (Kinging & Belz, 2005; van Compernelle, 2011);
- Hear-say evidential markers (Narita, 2012), interactional discourse markers (Yoshimi, 2001), and functions of *sumimasen* (Tateyama, 2001) in Japanese;
- Gambits in Spanish (Taylor, 2002);
- Modal particles in German (Belz & Vyatkina, 2005).

Key Concepts

Speech Event: Speech events include events accomplished largely by talk that may include multiple speech acts such as academic advising sessions, interviews, doctor's appointments, and service encounters.

Speech Act: Speech act theory views utterances not just as stating propositions, but as a way of doing things with words; hence the concept of *act* (Searle, 1969, 1976). Speech acts include five categories: asserting and explaining are representatives, requesting and advising are directives, promising and threatening are commissives, apologizing and complimenting are expressives, and declaring war and hiring/firing someone from a job are declaratives.

Semantic Formula: Sometimes also known as *pragmatic strategies*, semantic formulas are the component parts of a speech act. A speech act set specifies all the components of a given speech act, but most realizations of a speech act include a subset of possible semantic formulas. For example, an apology may include the head act "I'm sorry," and explanation "I didn't see you," a pledge of forbearance, "I won't happen again," or offer of repair, "I'll pay for the cleaning."

Kathleen Bardovi-Harlig

Pragmatics is a relative latecomer to the study of ISLA when compared to the other aspects of language (namely grammar, vocabulary, pronunciation, fluency, listening, and reading and writing), but in the relatively short history of the field of pragmatics and language learning, interest in instruction has been robust. The first survey of pragmatics research to address instruction in pragmatics was Kasper and Schmidt (1996), who summed up the research in instructed L2 pragmatics to date in one question: “Does instruction make a difference?” At the time, only six articles had been published, three that we would now call studies of effects of instruction (Billmyer, 1990; Bouton, 1994; Wildner-Bassett, 1984), and three additional assessments of language textbooks (Bardovi-Harlig, Hartford, Mahan-Taylor, Morgan, & Reynolds, 1991; Kasper, 1982; Scotton & Bernsten, 1988). Rose (2005) is arguably the seminal article on instructed L2 pragmatics. Rose (2005, p. 239; see also Kasper & Rose, 2002) posed three main questions, each of which has its own research methodology. Studies addressing the first question—“Is the targeted pragmatic feature teachable at all?”—employ pretest–posttest designs with intervening treatments. These studies show that pragmatic features can be learned from instruction, but they do not test the possibility that learners at the same proficiency could make equivalent progress without instruction, which forms Rose’s second question: “Is instruction in the targeted feature more effective than no instruction?” Studies addressing this question compare a control group that receives no pragmatics instruction to the treatment group. Studies of this type suggest that instruction has an advantage, but leave open the question of whether another type of intervention would have produced different outcomes. Thus, studies addressing the third question—“Are different teaching approaches differentially effective?”—compare two or more interventions and may include a control group with no pragmatics instruction.

The number of instructional effect studies has grown since Rose (2005) reviewed 25 articles (from 1986 to before 2005) that he called a “small, but growing body of research” (p. 386); Jeon and Kaya (2006) identified 34 studies (including unpublished studies) for their meta-analysis; and Takahashi (2010) reviewed 49, double the number reviewed by Rose (2005) only 5 years before. Taguchi (2015) reviewed 58 studies, selected to meet certain criteria, addressing two main questions: Is instruction effective in learning pragmatics? (cf. Rose’s Question 1) and What methods are most effective in learning pragmatics? (cf. Rose’s Question 3). The same year Bardovi-Harlig (2015b) reviewed 81 studies published between 2000 and mid-2013 to determine how conversation was operationalized in pragmatics instruction. These reviews show an increase in the number of investigations into the teaching of pragmatics, and an abiding interest in them.

The goals of instructional effect studies in pragmatics appear to be twofold: to determine what means of instruction facilitate or enhance the acquisition of L2 pragmatics through instruction, and to promote the teaching of pragmatics in second and foreign language classrooms worldwide. The former is consistent with the goals of ISLA. The second reflects the passion of many researchers, teachers, and researcher–teacher teams who are engaged in the study of instruction of pragmatics, with the goal of promoting research-based pragmatics teaching by demonstrating the efficacy of instruction and modeling how pragmatics can be taught. A single study may meet both goals.

Current Issues

One of the ongoing issues in studying pragmatics in ISLA is the lack of a pragmatics curriculum (for any language). Closely related to that is the lack of reference works (in any language) that catalogue the basic pragmatic phenomena for that language. It is hard to

imagine the teaching of grammar, pronunciation, fluency, vocabulary, listening, and reading and writing without the established associated pedagogies and reference works. Yet, in pragmatics there is no established approach that forms the basis for further inquiry.

Thus, this section considers the challenges for pragmatics instruction identified by Sykes (2013). These challenges are as relevant to research on the ISLA of pragmatics as to instruction itself, because every study of instructional effects has a pedagogical core around which it is built. Following Sykes, the eight challenges to teaching pragmatics are: (1) limited theoretical support for curricular development, (2) lack of authentic input in teaching materials, (3) lack of instructor knowledge, (4) a dominant focus on micro-features of language in the foreign language context, (5) time limitations in the classroom, (6) individual student differences and learning subjectivity, (7) feedback and assessment challenges, and (8) immense dialectal variation (Sykes, 2013, p. 73). To this list, I add (9) the lack of reference books and resources. In what follows, I briefly consider each of the issues raised by Sykes with the hope of encouraging researchers and teachers to work in these areas.

Limited Theoretical Support for Curricular Development

The research that has been conducted on acquisition of L2 pragmatics has revealed areas of difficulty for learners and can be taken as a needs assessment and a mandate for teaching L2 pragmatics (cf. Bardovi-Harlig, 2001). In the aggregate, L2 pragmatics research provides the content for L2 pragmatics instruction, although it may be noted that there is still no complete pragmatics curriculum. However, neither the specification of the target nor the development of language teaching pedagogy outside pragmatics is sufficient to fully specify a pedagogy of pragmatics, and as Kasper (2001) observed “it is not always obvious how principles proposed for instruction in grammar might translate to pragmatics” (p. 51). Kasper illustrates this claim by considering FonF, focus on form. When inappropriate utterances arise exclusively from their pragmalinguistics, that is, from the use of an inappropriate form, “a wrong discourse marker, routine formula, or modal verb to index illocutionary force or mitigation, for instance—and [is] limited to short utterance segments” (2001, p. 51), it may be possible to provide a recast. However, when sociopragmatics is involved, the source of inappropriate utterances or a sequence of utterances can depend on the context, a speaker’s interpretation of the situation or of another speaker’s turn, or a culturally determined assessment of what speech act is required given the situation, among many others. The fact that pragmatics is dependent on both the context and other speakers makes providing feedback challenging in any framework (this is discussed again in subsequent sections). The complexity of pragmatics also aligns more easily with a dichotomy of providing metapragmatic information (or not) rather than describing pragmatics instruction as explicit or implicit (see Kasper, 2001; Taguchi, 2015 for further discussion).

Research-based concerns, like the emphasis on authenticity in input and activities, has influenced the development of pragmatics pedagogy. Although researchers have made recommendations on what to teach, fewer proposals have been made on how to teach. Félix-Brasdefer and Cohen (2012) review a variety of models for teaching pragmatics, some of which specify content while others specify steps. Ohlstein and Cohen (1991) laid out five steps: conducting diagnostic assessment, presenting model dialogues, evaluating the situation, providing role play activities, and giving feedback and discussion. Martínez-Flor and Usó-Juan (2006) proposed six steps for teaching pragmatics: researching, reflecting, receiving, reasoning, rehearsing, and revising (“reviewing” in American English). Félix-Brasdefer (2006) suggested including three components focusing on content: communicative actions

Kathleen Bardovi-Harlig

and cross-cultural awareness, conversational analysis in the classroom, and communication practice. Koike's (2008) three-principle pedagogical model emphasizes content: contextualizing L2 grammar of pragmatics in natural context; providing grammatical, pragmatic, and sociocultural knowledge; and developing knowledge of sociopragmatic variation. Félix-Brasdefer and Cohen (2012) conclude their review by proposing four sequential components to teaching pragmatics: raising awareness, providing pragmatic input, teaching grammar as a communicative resource, and facilitating producing or production practice.

Lack of Authentic Pragmatic Input in Teaching Materials

The lack of authentic pragmatic input in commercially available second and foreign language textbooks has been well documented (Cohen & Ishihara, 2013; Eisenchlas, 2011; Vellenga, 2004). Reviews have shown that textbooks for both English as a Second Language (ESL) and English as a Foreign Language (EFL) present author-created conversations that do not reflect pragmatic usage by native speakers. The reviews have compared textbook presentations to natural or naturalistic conversations for a number of speech acts and pragmatic constructs, including but not limited to conversation closings (Bardovi-Harlig et al., 1991); pragmatic routines for agreement, disagreement, and clarifications (Bardovi-Harlig et al., 2015b); the social use of complaints (Boxer & Pickering, 1995); the language of business meetings (Williams, 1988); repair sequences (Cheng & Cheng, 2010); and, more generally, politeness (Limberg, 2016). The general state of pragmatics in current commercially marketed materials has led Cohen and Ishihara (2013, p. 116) to observe that "the actual dialogues may sound awkward or stilted, and are inauthentic in that they do not represent spontaneous pragmatic language as used in natural conversation."

Although textbooks do not meet the needs of pragmatics instruction for authenticity, there is a growing list of resources that teachers can use to teach or prepare pragmatics materials, including a book-length treatment, *Workplace Talk in Action: An ESOL Resource* (Riddiford & Newton, 2010). Additional resources include lessons developed by teachers in Bardovi-Harlig and Mahan-Taylor (2003), Tatsuki and Houck (2010), and Houck and Tatsuki (2011), and in Spanish, research-based website resources created by Félix-Brasdefer on the teaching of refusals (http://www.indiana.edu/~discprag/www_new/spch_refusals.html). Many researchers have also developed teaching activities, materials, and assessments, although these are often not available in published form. Nevertheless, without textbook presence, pragmatics will continue to be relegated to a supplemental rather than central status in the foreign- and second-language curriculum.

Lack of Instructor Knowledge

With respect to pragmatics, instructor knowledge involves knowing about pragmatics and knowing how to teach pragmatics. At the level of pragmatic knowledge, teachers may be familiar with the components of common speech acts as well as their functions, a range of conventional expressions, and regional pragmatic variation. Knowledge of teaching of pragmatics entails knowledge of pragmatics, but as suggested in a previous section, knowledge of pragmatics does not guarantee knowledge of how to teach it, as demonstrated by the fact that pragmatics pedagogy is still developing. Kasper (1997) laid out the issues and benefits of educating teachers in pragmatics in "The role of pragmatics in language teacher education." Ishihara and Cohen (2010) provide a book-length guide to the teaching of pragmatics for both novice teachers and experienced instructors already teaching pragmatics.

The consequence of lack of teacher knowledge in pragmatics for research on ISLA is that researchers either do their own teaching, which jeopardizes the impartiality of instruction, or they have to train teachers to do it (Bardovi-Harlig et al., 2015b; Eslami & Liu, 2013; Koike & Pearson, 2005).

Dominant Focus on Micro-features of Language in the Foreign Language Context

Grammar-dominated classes, whether in second or foreign language contexts, focus on micro-features. Advocates of pragmatics have allies in communicative, task-based, and content-based language teaching in moving classrooms away from micro-features. It may be possible to encourage teachers with a micro-feature orientation to explain how grammatical features are used pragmatically. For teachers and programs that organize the syllabus by micro-features, using Félix-Brasdefer and Cohen's (2012) approach to teaching grammar as a resource for pragmatics might be a way to begin. Félix-Brasdefer and Cohen provide a table that lists speech acts, associated grammar in Spanish, pragmatic function, and examples. Following the "grammar" column, a grammar-oriented lesson may illustrate the pragmatic uses of the subjunctive in Spanish by noting that its function is uncertainty, and that it occurs in speech acts of advice and suggestions. Similarly, a lesson on the conditional would note that its function is politeness and that it occurs in requests, refusals and disagreements, and asking for and giving directions.

Time Limitations in the Classroom

Time limitations are always an issue, even in intensive language programs. If we can advance the agenda of teaching pragmatics not as separate "special units" as in the case of almost every study cited in this chapter, but integrated into the main curriculum (and textbooks), it will cease to seem to be "extra" or "tacked on," competing for time with more traditional pedagogical targets.

Individual Student Differences and Learning Subjectivity

Individual learner differences and learner subjectivity capture how learners may vary. Learner differences capture observable characteristics such as age, gender, and proficiency, whereas learner subjectivity involves learners' own subjective categories including attitudes, perceptions of situations, and affective disposition, and is linked to social identity (Ishihara & Tarone, 2009; Seigal, 1996). Ishihara and Tarone (2009) distinguish between *accommodation* and *resistance*, which refer to learners' intended adoption or rejection of perceived L2 norms of which they are aware and linguistically capable of producing, and *convergence* and *divergence*, which refer to actual language use produced as a result of their accommodation or resistance. Because the use of L2 pragmatic norms involves speakers' social identity as well as their linguistic competence, it is not surprising that the literature reports learner resistance to adopting some L2 pragmatic norms. Reports of resistance include Western women's reluctance to adopt the high voice used by Japanese women or high levels of deference ("I don't want to be that humble"; Seigal, 1996) and Korean learners' rejection of Australian pragmatic routines when studying in Australia ("I rarely use Australian phrases when I speak English. It's because I feel uncomfortable. Australian English doesn't feel like my English. I mean, it feels unnatural to use some Australian phrases like 'Ta,' 'Good day mate,' 'No

worries' . . . stuff like that"; Davis, 2007, p. 629). In studies that investigate learner subjectivity (e.g., Eslami, Kim, Wright, & Burlbaw, 2014; Ishihara & Tarone, 2009; LoCastro, 2001; Seigal, 1996), learners are found to be positive overall to L2 pragmatics norms, but they also have limits that bar wholesale adoption of the target pragmatics. Acknowledging learner subjectivity, and a learner's right to it, poses interesting limitations on the assessment of pragmatic knowledge, which I consider briefly in the next section.

Feedback and Assessment Challenges

Feedback and evaluation in pragmatics can be challenging given that pragmatics is defined by choice: speakers make choices among available linguistic forms to convey social meanings. Because pragmatic value is derived from the *choice* of available linguistic devices to signal relationships among speakers, the study of acquisition of form in pragmatics—including grammar, lexicon, and formulaic language—is the study of the development of alternatives. The study of use in pragmatics must be understood in light of the forms available to the learner at any given stage of interlanguage development. There are many examples in the literature of choice between address terms (*Sie* versus *du*; Belz & Kinginger, 2003), request strategies (*would you* versus *I was wondering if you would*; Takahashi, 2005), or an aggravator rather than a mitigator (*I just decided on taking* versus *I was thinking about taking* or *I would like to take*; Bardovi-Harlig & Hartford, 1993). There is rarely one right answer but rather a range of felicitous alternatives. This contrasts sharply with grammaticality.

As in other areas of ISLA, feedback is postevent, or reactive (in contrast to models that are pre-event), occurring after learners have engaged in a production or interpretation activity, and may assume a variety of formats. Feedback has not been investigated to the same extent in pragmatics as it has been in other areas of language teaching, possibly due to the inherent challenge of multiple appropriate utterances in any given context.

Takenoya (2003) identifies an additional difficulty from the teacher's perspective, noting that teachers are sometimes uncomfortable making corrections. She observes that teachers of Japanese may feel that they are forcing American learners to behave like Japanese, or that they themselves are acting like mothers who are teaching manners to young children, roles that fit neither participant. Similarly, Thomas's (1983) early assessment of correction emphasizes the social aspect, adding the learners' perspective to the teachers':

Correcting pragmatic failure stemming from sociopragmatic miscalculation is a far more delicate matter for the language teacher than correcting pragmalinguistic failure. Sociopragmatic decisions are *social* before they are linguistic, and while foreign learners are fairly amenable to corrections which they regard as linguistic, they are justifiably sensitive about having their social (or even political, religious, or moral) judgment called into question.

p. 104, emphasis in original

If feedback is held to be pedagogically valuable, these issues must be addressed.

Assessment should match the goals of instruction and be consistent with pragmatics research. Accomplishing a speech act has many facets, and narrowing down what to measure can be an issue. Instructional effect studies sit right on the border of instruction, which tends to value "right" or "wrong" answers, and acquisition, which values developmental sequences and interlanguage forms. This distinction brings a certain tension to the scoring of learner production. Three fundamental principles of scoring are helpful when assessing the

influence of instruction on the linguistic system through production data (Bardovi-Harlig, 2015a):

1. Do not score so generously at the pretest that there is no room for improvement at the posttest (in a warning, for example, *be carefully* is not *be careful*).
2. Do not be so strict at the posttest that the analysis does not reveal improvement. The result will be that student responses were not “right” before instruction and still not “right” after instruction. For example, before instruction in a reciprocal thanking scenario when the organizer of a party thanks a student for coming (“Thank you for coming”), the student says “You’re welcome,” although acceptance of thanks is not the expected speech act. After instruction, some students say “Thank you for inviting me” and others say “Thank you for inviting Ø” (Bardovi-Harlig & Vellenga, 2012). The first response shows an adjustment of speech act (from acceptance of gratitude to reciprocal thanking) and the target-language conventional expression and the second response shows a change of speech act and the lexical core of the conventional expression, both of which show progress.
3. Take development into account. Use an interlanguage analysis by documenting what learners do. In the preceding example, we might use the interlanguage categories *speech act*, *lexical core of the conventional expression*, and *conventional expression*. The first example shows development in choice of speech act and use of conventional expression; the second shows development in choice of speech act and the lexical core of the conventional expression. Another example is found in the longitudinal study of academic advising sessions (Bardovi-Harlig, & Hartford, 1993). Advanced nonnative speakers showed changes over time in speech acts, semantic formulas, content, and form. Shifts in the speech acts performed (suggestions rather than refusals) were the first development observed, followed by semantic formulas. Content and form lagged behind. Even when the speech acts were performed with highly desirable mitigators, dispreferred aggravators sometimes also occurred in the same suggestion. Each component reflects a different level of knowledge and should be evaluated separately to give a fuller picture of development.

Key Concept

Interlanguage analysis: An interlanguage analysis provides an analysis of learner language as an independent system. Rather than evaluating an utterance as “right” or “wrong,” an interlanguage analysis attempts to describe learner production. There are often many developmental stages between “wrong” (the assumed starting point) and “right” (the targeted endpoint), and every step along the way represents progress.

The issues of scoring arise when researchers attempt to present a learner’s progress quantitatively. Qualitative analyses offer an alternative approach to assessment because they do not convert observations about linguistic development into scores (see for example, Couper, Denny & Watkins, 2015; Liddicoat & Crozet, 2001; Sardegna & Molle, 2010; Sydorenko & Tuason, in press). In these studies, changes in learner production are often presented in a series of examples.

Kathleen Bardovi-Harlig

With respect to learner subjectivity, Ishihara and Tarone (2009) suggest a solution to traditional assessment: learners have to understand the reason behind the L2 pragmatics and the social or pragmatic consequences of divergence. Resistant students may be given alternatives to performing the target, such as explaining what strategy they would use instead, which they then might be required to demonstrate during an assessment. Ishihara and Cohen (2010) provide other alternatives to traditional assessment: In one alternative (p. 305), a student may indicate her intention to make a request (1) the same way people in the community do, (2) more informally or more politely, or (3) not to use community norms. The student is then evaluated on how well she meets her own goal in (1) or (2) or how well she describes the norms that she does not want to follow and her reason in (3).

Dialectal Variation

Pragmatic variation was an open secret, discussed most by pragmaticists working on Spanish, until Barron and Schneider (2009) began to discuss variational pragmatics. A volume exploring variation in first and second language contexts appeared very soon after (Félix-Brasdefer & Koike, 2012). Barron and Schneider (2009) identified five main social variables relevant to pragmatic variation: region, social class, ethnicity, gender, and age. L2 pragmatics has long recognized variation due to age and gender, but there is also significant regional variation in languages, such as Chinese, English(es), and Spanish, whose speakers inhabit large or discontinuous geographical areas.

Pragmatics pedagogy has not yet sorted out its approach to regional variation, and there will, no doubt, be multiple solutions. One approach is to work with the local variety first. Working with conventional expressions and pragmatic routines, Bardovi-Harlig et al. (2015a, 2015b) identified pragmatic routines in the American Midwest (a dialect known as “General American”) for use by academic learners in the American Midwest from a corpus of academic English in the same region (namely the Michigan Corpus of Academic Spoken English, MICASE; Simpson, Briggs, Ovens, & Swales, 2002), thus matching both the academic environment and the regional variety.

Another approach is to give students a sampling from different regions and countries, starting with the local or the intended target area in foreign language teaching and expanding from there. The same cultural artifacts that promote dialect leveling in other linguistic domains (like television, movies, national news broadcasts, and plays) can be used pedagogically. By the same token, we can use the same to expose learners to other varieties. There are many ways to approach variation. The least promising is to ignore it.

Lack of Reference Books and Resources

To Sykes’s list, I add one additional challenge to the teaching of L2 pragmatics: the lack of reference books and resources. There is a connection between lack of reference materials, lack of authentic examples of conversation in textbooks, and lack of teacher knowledge. Over the years, my students have demonstrated this principle repeatedly. They are drawn to requests for both research and pedagogical development because the acquisition and use of requests by L2 speakers/learners have been investigated more than any other speech act. Requests are also the most likely speech act to be included in textbooks (often called “polite requests”), and are also the most researched speech act in instructional effect studies. The existence of resources encourages teaching in pragmatics; the lack of resources discourages it. If teachers do not have descriptions to draw on, they are unlikely

to undertake the primary research themselves for the sake of having an accurate description for teaching.

Reference books in pragmatics would not only provide access to materials on which teachers could base lessons, but would also assure the accuracy of pragmatic information. Explicit instruction requires accurate metapragmatic information, which even native speakers do not possess without education (Ishihara & Cohen, 2010; Wolfson, 1989), placing native and nonnative speaking teachers on the same footing. The closest thing to a reference work that has pragmatic information on multiple languages is the extensive pragmatics section of the CARLA website (Center for Advanced Research on Language Acquisition) hosted by the University of Minnesota (<http://www.carla.umn.edu/>). However, none of these resources compares to the multiple competing grammar reference works available from numerous publishers.

While developing a pedagogy of pragmatics faces several challenges, the empirical evidence that instruction can and does facilitate the acquisition of pragmatics indicates that meeting the challenges of developing such a pedagogy for use beyond an interested community of teachers and researchers is worthwhile.

Empirical Evidence

What we know from empirical evidence is that pragmatics can be learned through instruction. Rose's (2005) first two questions—"Is the targeted pragmatic feature teachable at all?" and "Is instruction in the targeted feature more effective than no instruction?" are answered affirmatively by all reviews. Pragmatics is teachable, and instruction surpasses no instruction. Rose's Question 3, "Are different teaching approaches differentially effective?" and Taguchi's (2015) Question 2, "What methods are most effective in learning pragmatics?" are much harder to answer. The key is that these questions bundle teaching as approaches and methods, rather than exploring them as multiple features. Taguchi recognizes this, breaking the comparisons of "implicit" and "explicit" (which dominate the field) as well as other labels for approaches to instruction into six main features rated on a binary scale for presence or absence: input enhancement, metapragmatic information, production, consciousness raising, feedback, and discussion. (She uses a seventh feature, input, which is present for all studies and thus not a distinguishing feature, although see Bardovi-Harlig, 2015b for a review of types of input used in pragmatics instruction). The type of task and assessment (Bardovi-Harlig, 2013) and authenticity and mode in the operationalization of conversation in input, practice, and assessment should also be investigated (Bardovi-Harlig, 2015b).

Although Taguchi is on the right track in breaking methods and approaches into features, she may not have gone far enough, because the reporting in instructional studies in pragmatics even at the feature level is often inconsistent. Taguchi reports that two features, the provision of direct metapragmatic information and production practice stand out as particularly effective. This section considers metapragmatic information, production practice, and adds feedback because it has not been as frequent an experimental variable in pragmatics as it has been in other areas of instruction.

Metapragmatic Information

Metapragmatic statements provide learners with information about the form, use, distribution, or other characteristics of the pragmatic construct selected for instruction. Alcón Soler (2007) embedded metapragmatic statements into a search activity where students looked for

Kathleen Bardovi-Harlig

examples in a transcript of a TV adventure series and provided an example for each statement. Two of her examples are provided here.

- (A) Imperatives are used to ask people to do something when one of the interlocutors has a higher position or they know each other very well.

Example: In making his request O'Neill uses an order (*Daniel, shut up*), which shows that he knows Daniel well. (p. 239)

- (B) In making requests the less you know someone or the higher the position someone has, the more polite and formal you need to be.

Example: O'Neill has a higher position than Carter, so he uses a conditional tense to indicate more polite language (*if you're gonna go back and tell General Hammond, I would like to stay here and take a look at their fusion technology*). (p. 240)

Key Concept

Metapragmatic information: Metapragmatic statements provide learners with information about the form, use, distribution, or other characteristics of the pragmatic construct selected for instruction.

In another approach to supplying metapragmatic information, Koike and Pearson (2005) gave information sheets to learners. One sheet listed seven ways to make suggestions in Spanish with a scale running from more to less direct; the second gave five ways of responding to a suggestion with the same scale. Takimoto's (2006) metapragmatic information was linked to a specific exercise, "the appropriateness score here should be four or five because the request is very polite with the use of lexical/phrasal downgraders" (p. 606). Examples of metapragmatic information are offered in a minority of studies, however. Metapragmatic information is more often only described without examples, as in "participants read a paragraph written in English summarizing the target form-function-context mappings" (Li, 2013, p. 50).

Metapragmatic information refers to the content of the information provided to the students. It can be placed in any sequence in instruction. Metapragmatic information that is given after a student makes an error of some type may be described as feedback, but given before an activity it may be described as part of the input.

Production Practice

Whereas it is often difficult to ascertain what metapragmatic information was given to learners, reports are generally better about describing production activities. Bardovi-Harlig (2015b) found that a range of production activities was used in the classrooms in which the studies were conducted. Oral practice for conversational pragmatics included conversations with native-speakers (Holmes & Riddiford, 2010; Sykes, 2005; Winke & Teng, 2010; Yoshimi, 2001), role plays (Eslami & Eslami-Rasekh, 2008; Félix-Brasdefer, 2008; Fukuya & Martínez-Flor, 2008), games (Bardovi-Harlig et al., 2015b), mock-job interviews (Louw et al., 2010), oral peer feedback in simulated writing groups (Nguyen, 2013), and problem-solving activities (Németh & Kormos, 2001). I have called this alignment of mode

oral-for-oral. Practice and targets in the same mode also include written practice for written production including written CMC (Belz & Kinginger, 2003; Belz & Vyatkina, 2005) and practice emails (written-for-written). In contrast, some studies use written practice to stand in for oral production by using written discourse completion tasks (DCTs) alone (Cohen & Shively, 2007; Eslami & Liu, 2013; written-for-oral) or with a mix of role plays and DCTs (Safont Jordá, 2003). Although highly desirable (Bardovi-Harlig, 2015a), the matched mode (oral-for-oral) in the form of role plays, conversations, and other oral practice activities, occurs in a higher proportion in practice tasks (among studies that include practice), than in assessment tasks (Bardovi-Harlig, 2015b). This may be because teachers do not have to assess the production activities in the same way that they have to evaluate the pretests and posttests and therefore the activities are more open-ended. In addition to the benefits of pushed output (Swain & Lapkin, 1995), trying out pragmatic acts in the L2 in different situations in the protected environment of the classroom may allow learners to experiment and explore using L2 pragmatics. When oral production activities are used, learners benefit from oral practice.

Key Concepts

Written Discourse Completion Task (DCT): Written DCTs are written production questionnaires that provide scenarios that typically include information about speakers and the context to which participants respond in writing as though they were speaking, as in this example from Eslami and Liu (2013, p. 71).

Your friend's birthday is coming and you are shopping for him/her. You see something in a display case that is appropriate as a gift. You want to look at it more closely. What would you say to the salesperson?

You: _____

Oral-for-oral: In oral production for oral production, the mode of elicited production aligns with the mode of production of the construct under investigation. To study conversation, for example, oral production may include natural conversation, role plays, and oral DCTs.

Written-for-written: In written production for written production, the mode of elicited production aligns with the mode of production of the construct under investigation. To study written communication, data may include natural communication including letters, notes, and written DCTs incorporating written scenarios such as sending an email.

Written-for-oral: When written production is used to study attributes of oral production, there is a mismatch in mode.

Feedback

Feedback as a variable has received very little attention in published studies of instructed pragmatics to date. Out of 58 studies, Taguchi (2015) lists 29 studies as using feedback. Of 81 studies reviewed in Bardovi-Harlig (2015b), I counted 34 with feedback. Of those only three studies investigated feedback as a variable (Barekat, 2013, who uses Takimoto's 2006 feedback script; Koike & Pearson, 2005; Takimoto, 2006), whereas the others simply reported it as a feature of the instruction. Takimoto provided feedback to learners after

Kathleen Bardovi-Harlig

they had made an incorrect selection from two possibilities in a written dual-choice task. Upon making an error in selection between two choices, learners received “immediate and explicit feedback on the correctness of the participants’ responses (i.e., “Can you find your error in your judgment? No, the appropriateness score here should be four or five because the request is very polite with the use of lexical/phrasal downgraders. Now next”) (2006, p. 606).

Koike and Pearson (2005) provided feedback after learners completed a series of exercises and activities. There were a control group and four experimental groups, two explicit and two implicit instruction groups. One explicit instruction group received explicit feedback, the other received implicit feedback; one implicit instruction group received explicit feedback, the other received implicit feedback. Koike and Pearson (2005) describe the feedback as follows:

For explicit feedback, learners were provided the correct answer after they presented their responses, and also some comment to reinforce why that answer was the most appropriate. For implicit feedback, learners were informed only whether their answer was correct by the teacher stating “Sí” ‘Yes’ or simply nodding or moving on to the next item, or incorrect by the teacher saying, “¿Cómo?” ‘What was that?’ or “Mm—no entendí” ‘Mm—I didn’t understand.’

p. 487

Koike and Pearson met with the participating instructors to give them specific instructions about the feedback strategies to be used for their classes in order to ensure that the classroom lessons matched the protocol established for the four treatment groups. As we can see, feedback was preset for the implicit condition, and had an established correct-answer-plus-reason format for the explicit condition. Koike and Pearson reported that explicit instruction with explicit feedback helped learners read, interpret, and select the most appropriate pragmatic choices in the multiple choice sections of the test, and the implicit feedback and possibly the implicit instruction led to an effect in open-ended responses in the dialogic context. In contrast, Takimoto reported no effect for feedback. In this case, both the type of feedback, and the type of activity were different. This suggests more focused investigation of the effect of feedback in pragmatics instruction and additional consideration of whether feedback interacts with the complexity of the language being corrected or with the instruction that precedes it.

Beyond these three studies, 31 studies employed feedback as part of instruction, but did not compare it to a nonfeedback or other-feedback condition. The presence of feedback in instructional designs for L2 pragmatics suggests to me that lesson designers view it as an integral part of instruction, even in the absence of studies that isolate feedback as a variable for investigation in instructed pragmatics.

In pragmatics instruction reported in the ISLA literature, feedback can vary in timing (immediate or delayed) and addressee (given to groups, individuals, or impersonally), and is variously operationalized as giving learners answer keys, which they use to compare to their own answers (Alcón Soler, 2007), group discussions summarizing issues especially in performed role plays (Félix-Brasdefer, 2008; Silva, 2003), and individualized feedback from the teacher, sometimes immediate and face-to-face (Takimoto, 2006), but also sometimes delayed face-to-face (Tateyama, 2007) or delayed by email (Ifantidou, 2013). Note that some of the activities used as feedback can also be found in other phases of instruction: comparing is often an activity found in implicit instruction, discussion often precedes production as

part of input or awareness-raising, and metapragmatic information may accompany input, as noted earlier.

Some instruction in pragmatics has utilized native speakers other than the teacher to give feedback. In Belz and Kinginger (2003) and Belz and Vyatkina (2005), native German-speaking keypals gave American learners of German spontaneous, immediate, and direct information on the use of familiar and formal address terms during telecollaboration. In Riddiford (2007) native speakers gave feedback on simulated job interviews in which the native speaker played the interviewers.

With respect to the most difficult aspect of feedback in pragmatics instruction, namely, appearing to correct someone's behavior rather than their language, and thus creating discomfort for both students and teachers, Holden and Sykes (2013) offer a potential solution through games. In one online game they developed, *Mentira*, students were assigned to a family. Each family had specific values and speech characteristics (such as being direct). When students selected speech that was inappropriate for the context (such as informal direct speech to a high-status character who is part of the game), they could be reprimanded or corrected.

Through these various contextualized interactions, learners see the impact of their pragmatic choices by learning to select behaviors relevant to each specific interaction or character. As a result, the same semantic formula has the effect of being extremely rude in one case and perfectly appropriate in another. A player is made aware of the success (or failure) of these choices through the NPCs' [nonplayer characters'] reactions and the assets that are awarded or taken away in the game (e.g., clues).

Holden & Sykes, 2013, p. 171

The immediate individualized feedback within the game—coupled with the option that a player may restart the game to test different speech acts and outcomes—may eliminate affective issues associated with being corrected and is highly promising. Although not everyone can program their own online pragmatics games, these principles can be used in board games with card draws (e.g., Bardovi-Harlig et al., 2015a, 2015b).

Pedagogical Implications

The findings that pragmatics can be learned from instruction and that instruction is superior to no instruction warrant continued efforts to move pragmatics teaching into the mainstream. Given the state of affairs described by Sykes (2013), one of the clearest ways to ensure the inclusion of pragmatics in foreign and second language curricula is to educate language teachers, strengthening both their knowledge about pragmatics and pragmatics instruction. Following teacher education, other attainable pedagogical goals include the improvement of pragmatic representations in textbooks and reference works, and the subsequent integration of pragmatics into language teaching.

To that end, practitioners, researchers, and teacher-educators can work together to promote the development of teaching materials and activities for pragmatics in a variety of ways. One place to begin is with the collection of authentic examples. One of the most important elements in the teaching of pragmatics is to provide learners with authentic models for input, practice, and assessment. Teachers can collect examples as they occur and gradually build up a usable file and create a shared repository with teachers in the same program, school, or professional association. Inappropriate examples are relatively easier to collect because they

Kathleen Bardovi-Harlig

draw our attention, but appropriate examples are important too. An especially nice thank-you note or engaging invitation might also go into the file. Recording how messages made one feel may enhance further discussion. Email, texts, posts, and voice mail make it particularly easy to collect examples. Face-to-face conversations or phone conversations can be jotted down or dictated to a device, and rerecorded for use in the classroom.

When developing input, practice, and production in pragmatics instruction, matching oral-for-oral is highly desirable. Most instruction in pragmatics is targeted at conversational use of language. Learners need to be able to hear (and see) input and to say their turns in conversation. That is not to say that steps to learning cannot incorporate written information, such as transcripts and metapragmatic information, but rather that all instructional units that have conversation as the ultimate target should also contain oral production as well as aural input. The same point should be made to practice written communication in writing, but there are no examples of mismatched mode when writing is the target.

Free online corpora can help teachers build repertoires of authentic and authentic-scripted language. Teachers should choose a corpus that matches the target: for example, spoken academic English can be found in the Michigan Corpus of Academic Spoken English (MICASE; Simpson et al., 2002; <http://quod.lib.umich.edu/m/micase>); conversation among family and friends can be found in the Santa Barbara Corpus of Spoken American English (Du Bois Chafe, Meyer, Thompson, Englebretson, & Martey, 2000–2005; <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>); and television interviews and talk shows can be found in the Corpus of Contemporary American English (COCA; Davies, 2008; <http://corpus.byu.edu/coca/>). Corpora are often easy enough to search that advanced learners can also use them. Information on working with corpora to develop pragmatics units and lessons can be found in Bardovi-Harlig and Mossman (2016) and Ishihara and Cohen (2010). Another type of corpus found on the internet are fan transcriptions of popular television shows. Fifty thousand words for each of five dramas and five sitcoms were collected and organized with a concordance tool on the Compleat Lexical Tutor by Tom Cobb under the heading “TV-Marlise” (<http://www.lexutor.ca/conc/eng>), but anyone interested in a particular television show could locate it by using an online search engine. Transcripts can often be paired with the broadcast version, which provides the audio and visual cues that are necessary for conversation and learning conversational pragmatics.

Learning to work with textbooks as a starting point is an important step in teaching pragmatics. Although current commercially available textbooks do not provide learners with authentic pragmatic input, they can be used as a starting point for classroom activities. For example, textbook conversations often lack closings or include partial closings. To illustrate the importance of closings, we asked students to read a conversation from the textbook, then walk away when they reached the end (Bardovi-Harlig et al., 1991). The students immediately recognize what happened, and this is a very effective consciousness-raising activity. Another way to work with the textbooks is to check the “useful expressions” provided by textbooks against a corpus; advanced learners can also do this. They will find that some expressions are truly useful, that some have very low occurrence, and that others do not occur at all. They will also learn that some expressions do not occur in some corpora, but occur in others.

Teachers should be ready to engage in discovery activities with their students. As noted earlier, with the absence of compiled reference works, it is often up to teachers and students to make new discoveries. Exploring pragmatics with the students works at all levels whether students are aspiring teachers or language learners. Teachers can devise projects for teams where students work together. In teacher preparation courses, we often begin by

having students investigate a speech act that is well researched. I then have students choose a speech act that is not listed on the CARLA website or has only a few citations. Students can variously elect to investigate a pragmatics problem that bothers them (native speakers and nonnative speakers alike have noticed pragmatic usages that are worth pursuing) or one in which they are interested for other reasons. The instructor does not have to be the expert as much as a guide through the investigation process. A valuable tangible outcome to the investigations by students is a “published” booklet or website that summarizes what was used and that includes participant insights.

A final recommendation is to create a pragmatics repository with other teachers. Having ready access to authentic materials for teaching enhances the development of teaching materials. A repository of teaching materials for pragmatics will encourage more integration of pragmatics into the curriculum. When teachers do not have to start at the beginning for every lesson, they can revise, elaborate, or enhance what has already been contributed. Local repositories have the advantage of being regionally appropriate and suited to the needs of students in particular programs. Cohen (2016) documents the development of websites that promote the teaching of L2 pragmatics. He has established a new site, Second and Foreign Language Pragmatics Wiki (<http://wlp pragmatics.pbworks.com>) as a repository for teaching materials in a variety of target languages. Although not a substitute for local collaboration within a program, such a repository can provide a much broader perspective and reach a larger range of teachers and researchers.

Teaching Tips

- Collect authentic examples, good as well as bad.
- When developing input, practice, and production, match oral-for-oral in pragmatics.
- Become acquainted with corpora.
- Learn to work with textbooks as a starting point.
- Engage in discovery activities with your students.
- Create a pragmatics repository with teachers in your program.

Future Directions

In this section I discuss three main future directions: promoting replication, isolating instructional variables, and studying the interaction of variables.

Replication

One clear future direction is to meet the goal of replicability. Studies can be replicated only when sufficient information is provided by the original research reports. Examples must be provided for all relevant categories, including input, metapragmatics, discussions, input enhancement, and feedback, not to mention examples of scoring of the language samples that are being investigated. For replications to be possible, enough information about the task and the teaching has to be provided. A related goal would be to establish an instructional materials archive that parallels the research archive (IRIS) for research data collection tasks.

Full specification of pedagogical treatment is beyond the scope (and word limit) of most published ISLA research studies on pragmatics, but there are some paired publications in which researchers report the research in one and the instruction in the other. Riddiford and

Kathleen Bardovi-Harlig

her colleagues documented their research on workplace requests and sociopragmatics (Riddiford, 2007; Holmes & Riddiford, 2010; Riddiford & Joe, 2010). Riddiford and Newton (2010) developed pedagogical materials. Nguyen and colleagues reported research on teaching constructive criticism (Nguyen, 2013; Nguyen, Pham, & Cao, 2013; Nguyen, Pham, & Pham, 2012) and discussed teaching constructive criticism (Nguyen & Basturkmen, 2010). Bardovi-Harlig, Mossman, and Vellenga tested the efficacy of teaching pragmatic routines of agreeing, disagreeing, and clarification for academic group work and then described the process of materials development for teachers (Bardovi-Harlig, Mossman, & Vellenga, 2015b, 2015a, respectively; see Bardovi-Harlig, 2015a for further elaboration).

Isolating Variables/Factors

In keeping with work in ISLA in other areas (see other chapters in this volume), we should work to separate variables that are often bundled in instruction. Implicit and explicit presentations of information, or metapragmatic information, can occur as instruction before learners engage in activities or following activities as feedback. Feedback can be given explicitly or implicitly during or after activities. Discussion may occur during instruction as part of pragmatic awareness-raising or as feedback in response to performance on an activity. Making a distinction between the type of information given or the type of activity engaged in and its sequencing, between what happens in instruction, and what happens in feedback, facilitates a clearer comparison.

Exploring Interaction of Variables/Factors

Once the instructional variables have been isolated, we can begin to systematically explore the relation between them and ultimately determine their effect on acquisition. In a previous review of types of input rated by authenticity and mode of delivery (Bardovi-Harlig, 2015b), I suggested exploring two additional questions in the ISLA of pragmatics: “How does the representation of conversation in input affect pragmatic learning?” and “How does the operationalization of conversation during practice activities affect pragmatic learning?” Gilmore (2011) began to address the first issue when he compared input from a textbook and authentic input in a study of L2 communicative competence over the course of a semester.

Here I suggest investigating the richness (and complexity) of input and the relative efficacy of different types of noticing activities, whether they include metapragmatic statements, guided noticing, input enhancement, or other instructional means of assisting learners toward awareness of target features. Are the benefits of one means of drawing attention limited when the input is extremely simple or reduced? Is it possible, for example, that any difference in feedback or presentation is lost when alternatives are limited as in the case of a binary-choice task? For example, in Takimoto’s study (2006), input consists of invented textbook-style written conversations, in which learners are asked to pick one of two choices as appropriate, then feedback on an incorrect selection tells them that the other choice has a better politeness marker. Given that there are only two choices, it is possible that learners could arrive at “not a, b” by themselves and do not require an explanation. The input itself may put a ceiling on how helpful either feedback or noticing can be. In contrast, learners who are trying to take in the multiple features of authentic or elicited conversations including refusals that take place across multiple turns with a variety of grammatical and lexical modifiers could arguably need more help. Because only a few studies investigate the effect of feedback directly, there is an opportunity to design new studies with attention to additional variables unique to pragmatics.

Conclusion

This is a very exciting time in the development of research in the ISLA of pragmatics. Much remains to be done, and there is both need and interest in the field. From pedagogical development, to teacher education, to better planned and better reported studies, there is a role for everyone who wants one.

References

- Alcón Soler, E. (2007). Fostering EFL learners' awareness of requesting through explicit and implicit consciousness-raising tasks. In M. del Pilar Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 221–241). Bristol: Multilingual Matters.
- Bardovi-Harlig, K. (2001). Evaluating the empirical evidence: Grounds for instruction in pragmatics? In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 13–32). Cambridge: Cambridge University Press.
- Bardovi-Harlig, K. (2010). Pragmatics and second language acquisition. In R. Kaplan (Ed.), *The handbook of applied linguistics* (2nd ed., pp. 232–243). Oxford: Oxford University Press.
- Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning*, 63 suppl. 1, 68–86.
- Bardovi-Harlig, K. (2015a). Designing instructional effect studies for L2 pragmatics: A guide for teachers and researchers. In S. Gesuato, F. Bianchi, & W. Cheng (Eds.), *Teaching, learning and investigating pragmatics: Principles, methods and practices* (pp. 135–164). Newcastle upon Tyne: Cambridge Scholars.
- Bardovi-Harlig, K. (2015b). Operationalizing conversation in studies of instructional effects in L2 pragmatics. *System*, 48, 21–34.
- Bardovi-Harlig, K., & Hartford, B. S. (1993). Learning the rules of academic talk: A longitudinal study of pragmatic development. *Studies in Second Language Acquisition*, 15, 279–304.
- Bardovi-Harlig, K., Hartford, B. A. S., Mahan-Taylor, R., Morgan, M. J., & Reynolds, D. W. (1991). Developing pragmatic awareness: Closing the conversation. *ELT Journal*, 45, 4–15.
- Bardovi-Harlig, K., & Mahan-Taylor, R. (2003). *Teaching pragmatics*. Washington, DC: United States Department of State. Retrieved from <http://americanenglish.state.gov/search/solr/pragmatics>
- Bardovi-Harlig, K., & Mossman, S. (2016). Corpus-based materials development for teaching and learning pragmatic routines. In B. Tomlinson (Ed.), *SLA research and materials development for language learning* (pp. 250–267). London: Taylor & Francis.
- Bardovi-Harlig, K., Mossman, S., & Vellenga, H. E. (2015a). Developing corpus-based materials to teach pragmatic routines. *TESOL Journal*, 6, 499–526.
- Bardovi-Harlig, K., Mossman, S., & Vellenga, H. E. (2015b). The effect of instruction on pragmatic routines in academic discussion. *Language Teaching Research*, 19, 324–350.
- Bardovi-Harlig, K., & Vellenga, H. E. (2012). The effect of instruction on conventional expressions in L2 pragmatics. *System*, 40, 77–89.
- Barekat, B. (2013). Investigating effects of metalinguistic feedback in L2 pragmatic instruction. *International Journal of Linguistics*, 5, 197–208.
- Barron, A., & Schneider, K. P. (2009). Variational pragmatics: Studying the impact of social factors on language use in interaction. *Intercultural Pragmatics*, 6, 425–442.
- Belz, J. A., & Kinginger, C. (2003). Discourse options and the development of pragmatic competence by classroom learners of German: The case of address forms. *Language Learning*, 53, 591–647.
- Belz, J. A., & Vyatkina, N. (2005). Learner corpus analysis and the development of L2 pragmatic competence in networked intercultural language study: The case of German modal particles. *Canadian Modern Language Review*, 62, 17–48.
- Billmyer, K. (1990). “I really like your lifestyle”: ESL learners learning how to compliment. *Penn Working Papers in Educational Linguistics*, 6, 31–48.
- Bouton, L. F. (1994). Can NNS skill in interpreting implicatures in American English be improved through explicit instruction? A pilot study. In L. F. Bouton & Y. Kachru (Eds.), *Pragmatics and*

Kathleen Bardovi-Harlig

- language learning* (Vol. 5, pp. 88–109). Urbana-Champaign: University of Illinois, Division of English as an International Language.
- Boxer, D., & Pickering, L. (1995). Problems in the presentation of speech acts in ELT materials: The case of complaints. *ELT Journal*, 49, 44–58.
- Cheng, W., & Cheng, P. (2010). Correcting others and self-correcting in business and professional discourse and textbooks. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (Vol. 7 of Handbooks of pragmatics, pp. 443–466). Berlin: Mouton de Gruyter.
- Cobb, T. Corpus concordance English [Computer Program]. Accessed November 16, 2014 @ <http://www.lexutor.ca/conc/eng>
- Cohen, A. D. (2016). The Design and construction of websites to promote L2 pragmatics. In K. Bardovi-Harlig, & J.C. Félix-Brasdefer (Eds.), *Pragmatics and Language Learning*, (Vol. 14, pp. 341–356). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Cohen, A. D., & Ishihara, N. (2013). Pragmatics. In B. Tomlinson (Ed.), *Applied linguistics and materials development* (pp. 113–126). London: Bloomsbury Academic.
- Cohen, A. D., & Shively, R. (2007). Acquisition of requests and apologies in Spanish and French: Impact of study abroad and strategy-building intervention. *Modern Language Journal*, 91, 189–212.
- Couper, G., Denny, H., & Watkins, A. (2015). Teaching the sociocultural norms of an undergraduate community of practice. *TESOL Journal*, 7, 4–39.
- Crystal, D. (Ed.). (1997). *The Cambridge encyclopedia of language* (2nd ed.). New York: Cambridge University Press.
- Davies, M. (2008). *The corpus of contemporary American English: 450 million words, 1990-present*. Retrieved from <http://corpus.byu.edu/coca/>
- Davis, J. (2007). Resistance to L2 pragmatics in the Australian ESL context. *Language Learning*, 57, 611–649.
- Du Bois, J. W., Chafe, W. L., Meyer, C., Thompson, S. A., Englebretson, R., & Martey, N. (2000–2005). *Santa Barbara corpus of spoken American English, parts 1–4*. Philadelphia: Linguistic Data Consortium. Retrieved from <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>
- Eisenchlas, S. A. (2011). On-line interactions as a resource to raise pragmatic awareness. *Journal of Pragmatics*, 43, 51–61.
- Eslami, Z. R., & Eslami-Rasekh, A. (2008). Enhancing the pragmatic competence of non-native English-speaking teacher candidates (NNESTCs) in an EFL context. In E. Alcón Soler & A. Martínez-Flor (Eds.), *Investigating pragmatics in foreign language learning, teaching and testing* (pp. 178–197). Bristol: Multilingual Matters.
- Eslami, Z. R., Kim, H., Wright, K. L., & Burlaw, L. M. (2014). The role of learner subjectivity and Korean-English language learners pragmatic choices. *Lodz Papers in Pragmatics*, 10, 117–146.
- Eslami, Z. R., & Liu, C. (2013). Learning pragmatics through computer-mediated communication in Taiwan. *International Journal of Science, Culture and Language*, 1, 52–73.
- Félix-Brasdefer, J. C. (2006). Teaching the negotiation of multi-turn speech acts: Using conversation-analytic tools to teach pragmatics in the classroom. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning* (Vol. 11, pp. 165–197). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Félix-Brasdefer, J. C. (2008). Teaching Spanish pragmatics in the classroom: Explicit instruction of mitigation. *Hispania*, 91, 477–492.
- Félix-Brasdefer, J. C., & Cohen, A. D. (2012). Teaching pragmatics in the foreign language classroom: Grammar as a communicative resource. *Hispania*, 95, 650–669.
- Félix-Brasdefer, J. C., & Koike, D. (Eds.). (2012). *Pragmatic variation in first and second language contexts: Methodological issues*. Amsterdam/Philadelphia: John Benjamins.
- Fukuya, Y., & Martínez-Flor, A. (2008). The interactive effects of pragmatic-eliciting tasks and pragmatic instruction. *Foreign Language Annals*, 41, 478–500.
- Gilmore, A. (2011). “I prefer not text”: Developing Japanese learners’ communicative competence with authentic materials. *Language Learning*, 61, 786–819.

- Holden, C. L., & Sykes, J. M. (2013). Complex L2 pragmatic feedback via place-based mobile games. In N. Taguchi & J. M. Sykes (Eds.), *Technology in interlanguage pragmatics research and teaching* (pp. 155–170). Philadelphia: John Benjamins.
- Holmes, J., & Riddiford, N. (2010). From classroom to workplace: Tracking socio-pragmatic development. *ELT Journal*, 65, 376–386.
- Houck, N. R., & Tatsuki, D. H. (Eds.). (2011). *Pragmatics: Teaching natural conversation*. New York: TESOL.
- Ifantidou, E. (2013). Pragmatic competence and explicit instruction. *Journal of Pragmatics*, 59, 93–116.
- Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. London: Longman.
- Ishihara, N., & Tarone, E. (2009). Subjectivity and pragmatic choice in L2 Japanese: Emulating and resisting pragmatic norms. In N. Taguchi (Ed.), *Pragmatic competence* (pp. 101–128). Berlin: Mouton de Gruyter.
- Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 165–211). Amsterdam: John Benjamins.
- Kasper, G. (1982). Teaching-induced aspects of interlanguage discourse. *Studies in Second Language Acquisition*, 4, 99–113.
- Kasper, G. (1997). The role of pragmatics in language teacher education. In K. Bardovi-Harlig & B. S. Hartford (Eds.), *Beyond methods: Components of language teacher education* (pp. 113–136). New York: McGraw Hill.
- Kasper, G. (2001). Classroom research on interlanguage pragmatics. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 33–60). Cambridge: Cambridge University Press.
- Kasper, G., & Rose, K. (2002). *Pragmatic development in a second language*. Malden, MA: Wiley-Blackwell.
- Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies in Second Language Acquisition*, 18, 149–169.
- Kinging, C., & Belz, J. (2005). Socio-cultural perspectives on pragmatic development in foreign language learning: Microgenetic case studies from telecollaboration and residence abroad. *Intercultural Pragmatics*, 2, 369–421.
- Koike, D. (2008). A grammar of L2 pragmatics: Issues in learning and teaching. In S. Katz & J. Watzinger-Tharp (Eds.), *Conceptions of L2 grammar: Theoretical approaches and their application in the L2 classroom* (pp. 35–52). Boston, MA: Heinle.
- Koike, D., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System*, 33, 481–501.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Li, S. (2013). Amount of practice and pragmatic development of request-making in L2 Chinese. In N. Taguchi & J. M. Sykes (Eds.), *Technology in interlanguage pragmatics research and teaching* (pp. 44–69). Philadelphia: John Benjamins.
- Liddicoat, A., & Crozet, C. (2001). Acquiring French interactional norms through instruction. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 125–144). Cambridge: Cambridge University Press.
- Limberg, H. (2016). “Always remember to say *please* and *thank you*”: Teaching politeness with German EFL textbooks. In K. Bardovi-Harlig & C. Félix-Brasdefer (Eds.), *Pragmatics and language learning* (Vol. 14, pp. 265–292). Honolulu: University of Hawai’i, National Foreign Language Resource Center.
- LoCastro, V. (2001). Individual differences in second language acquisition: Attitudes, learner subjectivity, and L2 pragmatic norms. *System*, 29, 69–89.
- Louw, K., Derwing, T., & Abbott, M. (2010). Teaching pragmatics to L2 learners for the workplace: The job interview. *Canadian Modern Language Review*, 66, 739–758.
- Martínez-Flor, A., & Usó-Juan, E. (2006). A comprehensive pedagogical framework to develop pragmatics in the foreign language classroom: The 6Rs approach. *Applied Language Learning*, 16, 39–64.

- Narita, R. (2012). The effects of pragmatic consciousness-raising activity on the development of pragmatic awareness and use of hearsay evidential markers for learners of Japanese as a foreign language. *Journal of Pragmatics*, 44, 1–29.
- Németh, N., & Kormos, J. (2001). Pragmatic aspects of task performance: The case of argumentation. *Language Teaching Research*, 5, 213–240.
- Nguyen, T.T.M. (2013). Instructional effects on the acquisition of modifiers in constructive criticism by EFL learners. *Language Awareness*, 22, 76–94.
- Nguyen, T.T.M., & Basturkmen, H. (2010). Teaching constructive critical feedback. In D. H. Tatsuki & N. R. Houck (Eds.), *Pragmatics: Teaching speech acts* (pp. 125–140). New York: TESOL.
- Nguyen, T.T.M., Pham, M. T., & Cao, T. H. (2013). The effects of explicit metapragmatic instruction on EFL learners' performance of constructive criticism in an academic setting. In T. Greer, D. Tatsuki & C. Roever (Eds.), *Pragmatics and language learning* (Vol. 13, pp. 213–244). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Nguyen, T.T.M., Pham, T. H., & Pham, M. T. (2012). The relative effects of explicit and implicit form-focused instruction on the development of L2 pragmatic competence. *Journal of Pragmatics*, 44, 416–434.
- Olshtain, E., & Cohen, A. D. (1991). Teaching speech act behavior to nonnative speakers. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 154–165). New York: Newbury House.
- Riddiford, N. (2007). Making requests appropriately in a second language: Does instruction help to develop pragmatic proficiency? *TESOLANZ Journal*, 15, 88–102.
- Riddiford, N., & Joe, A. (2010). Tracking the development of sociopragmatic skills. *TESOL Quarterly*, 44, 195–205.
- Riddiford, N., & Newton, J. (2010). *Workplace talk in action: An ESOL resource*. Wellington: School of Linguistics and Applied Language Studies, Victoria University of Wellington.
- Rose, K. R. (2005). On the effects of instruction in second language pragmatics. *System*, 33, 385–399.
- Safont Jordá, M. P. (2003). Instructional effects on the use of request acts modification devices by EFL learners. In A. Martínez-Flor, E. Usó-Juan, & A. Fernández-Guerra (Eds.), *Pragmatic competence and foreign language teaching* (pp. 211–232). Castelló de la Plana: Universitat Jaume I.
- Safont Jordá, M. P., & Alcón Soler, E. (2012). Teachability of request act peripheral modification devices in third language learning contexts. In M. Economidou-Kogetsidis & H. Woodfield (Eds.), *Interlanguage request modification* (pp. 275–313). Amsterdam, The Netherlands: John Benjamins.
- Sardegna, V., & Molle, D. (2010). Videoconferencing with strangers: Teaching Japanese EFL students verbal backchannel signals and reactive expressions. *Intercultural Pragmatics*, 7, 279–310.
- Scotton, C. M., & Bernstein, J. (1988). Natural conversations as a model for textbook dialogue. *Applied Linguistics*, 9, 372–384.
- Searle, J. R. (1969). *Speech acts*. Cambridge: Cambridge University Press.
- Searle, J. R. (1976). A classification of illocutionary acts. *Language and Society*, 5, 1–23.
- Seigal, M. (1996). The role of learner subjectivity in second language sociolinguistic competency: Western women learning Japanese. *Applied Linguistics*, 17, 356–382.
- Silva, A. (2003). The effects of instruction on pragmatic development: Teaching polite refusals in English. *Second Language Studies*, 22, 55–106.
- Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). The Michigan corpus of academic spoken English. Retrieved from <http://quod.lib.umich.edu/m/micase>
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371–391.
- Sydorenko, T., & Tuason, G. H. (2016). Noticing of pragmatic features during spoken interaction. In K. Bardovi-Harlig & C. Félix-Brasdefer (Eds.), *Pragmatics and language learning* (Vol. 14, pp. 233–264). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
- Sykes, J. M. (2005). Synchronous CMC and pragmatic development: Effects of oral and written chat. *CALICO Journal*, 22, 399–432.

- Sykes, J. M. (2013). Multiuser virtual environments: Learner apologies in Spanish. In N. Taguchi & J. M. Sykes (Eds.), *Technology in interlanguage pragmatics research and teaching* (pp. 71–100). Philadelphia: John Benjamins.
- Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching*, 48, 1–50.
- Takahashi, S. (2005). Pragmalinguistic awareness: Is it related to motivation and proficiency? *Applied Linguistics*, 26, 90–120.
- Takahashi, S. (2010). Assessing learnability in second language pragmatics. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (Handbooks of pragmatics, Vol. 7, pp. 391–421). Berlin: Mouton de Gruyter.
- Takenoya, M. (2003). *Terms of address in Japanese: An interlanguage pragmatics approach*. Sapporo: Hokkaido University Press.
- Takimoto, M. (2006). The effects of explicit feedback and form–meaning processing on the development of pragmatic proficiency in consciousness-raising tasks. *System*, 34, 601–614.
- Tateyama, Y. (2001). Explicit and implicit teaching of pragmatic routines. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 200–222). Cambridge: Cambridge University Press.
- Tateyama, Y. (2007). The effects of instruction on pragmatic awareness. In K. Bardford-Watts (Ed.), *JALT 2006 conference proceedings* (pp. 1189–1200). Tokyo: JALT.
- Tatsuki, D. H., & Houck, N. R. (Eds.). (2010). *Pragmatics: Teaching speech acts* (pp. 163–180). New York: TESOL.
- Taylor, G. (2002). Conversation strategies by intermediate Spanish students. *Foreign Language Annals*, 35, 171–189.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4, 91–111.
- Trosborg, A., & Shaw, P. (2008). Deductive and inductive methods in the teaching of business pragmatics: Not an either/or. In R. Geluykens & B. Kraft (Eds.), *Institutional discourse in cross-cultural contexts (LINCOM studies in Pragmatics 14)* (pp. 193–220). Munchen: LINCOM.
- van Compernelle, R. (2011). Developing second language sociopragmatic knowledge through concept-based instruction: A microgenetic case study. *Journal of Pragmatics*, 43, 3267–3283.
- Vellenga, H. E. (2004). Learning pragmatics from ESL and EFL textbooks: How likely? *TESL-EJ*, 8(2), 1–18.
- Wildner-Bassett, M. (1984). *Improving pragmatic aspects of learners' interlanguage*. Tübingen: Narr.
- Williams, M. (1988). Language taught for meetings and language used in meetings: Is there anything in common? *Applied Linguistics*, 9, 45–58.
- Winke, P., & Teng, C. (2010). Using task-based pragmatics tutorials while studying abroad in China. *Intercultural Pragmatics*, 7, 363–399.
- Wolfson, N. (1989). *Perspectives: Sociolinguistics and TESOL*. Rowley, MA: Newbury House.
- Yoshimi, D. R. (2001). Explicit instruction and JFL learners' use of interactional discourse markers. In K. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 223–244). Cambridge: Cambridge University Press.

L2 Fluency Development

Tracey M. Derwing

Background

In a classic paper, Charles Fillmore (2000, originally published in 1979) characterized four types of first language (L1) fluency, the simplest of which is the “ability to talk at length with few pauses” (p. 51) while the most complex is some people’s ability to “be creative and imaginative in their language use, to express ideas in novel ways, to pun, to make up jokes, to attend to the sound independently of the sense, to vary styles, to create and build on metaphors, and so on” (p. 51).

In second language (L2) studies, most research efforts have focused on some version of Fillmore’s first category; in many instances temporal measures of L2 productions are made, such as syllables per second, number and length of pauses, and mean length of run (mean number of syllables between pauses). Many of these studies are motivated by a pedagogical concern to determine the effects of speaking task type, topic, preplanning, and various instructional activities, with the goal of identifying ways to help L2 learners become more fluent or fluid in their speech. As several researchers have noted, the lay use of the term ‘fluency’ to mean ‘proficiency level’ as in the sentence “Murray is fluent in French” is not the intent of most applied linguistics studies. ‘Fluency’ in L2 productions, especially in speech research, generally refers to the degree to which speech flows, and to what extent that flow is interrupted by pauses, hesitations, false starts, and so on.

However, as Segalowitz (2010) has pointed out, from a psycholinguistic perspective, there are three types of fluency, the first of which is *cognitive fluency*: “the speaker’s ability to efficiently mobilize and integrate the underlying cognitive processes responsible for producing utterances with the characteristics that they have” (p. 48). In other words, factors such as short-term memory, planning, lexical retrieval, and appropriate choice of grammatical markers are involved, in addition to the suppression of the first language (L1) and other L2s. Segalowitz’s second definition is *utterance fluency*: “the temporal, pausing, hesitation and repair characteristics” (p. 48) of an utterance; these features of L2 speech are the oral manifestations of the speaker’s level of cognitive fluency (e.g., Derwing, Rossiter, Munro, & Thomson, 2004; Lennon, 1990, see Table 14.1). Finally, there

Table 14.1 Common measures of utterance fluency

Fluency Measure	Definition
Mean length of pauses	Average length of pauses in milliseconds
Pauses/second	Number of pauses per second
# of filled pauses/syll	filled pauses (fillers) are nonwords such as 'um,' 'er,' 'uh.' This measure usually involves counting how many pauses there are, divided by total number of syllables in the speech sample.
# of unfilled pauses/syll	The number of silent pauses divided by the total number of syllables in the speech sample
Mean length of run (MLR)	Average # of syllables between unfilled pauses
Speech rate	Syllables per second including pauses
Articulation rate	Syllables per second after removal of pauses
False starts	Abandoned portions of an utterance that are followed by a new approach
Self-repetitions	An exact repetition of a word or phrase
Pruned syllables/second	Remaining syllables after removing nonlexical filled pauses, self-corrections, false starts, self-repetitions and asides, divided by the number of seconds in the speech sample

is *perceived fluency*, judgments “made about speakers based on impressions drawn from their speech samples” (p. 48). Judgments (often made on a Likert scale ranging from ‘extremely fluent’ to ‘extremely dysfluent’) have been shown to be correlated with several aspects of utterance fluency (e.g., Derwing et al., 2004; Rossiter, 2009). In some ways these measures are also proxies for problems at the cognitive level; a memory problem or word finding difficulties will result in delays as a speaker tries to produce an intended meaning in real time.

Why is L2 fluency important? A major disadvantage of perceived dysfluency and limited utterance fluency is that listeners can find it tiring and annoying to attend to highly dysfluent speech (Varonis & Gass, 1982). If potential interlocutors avoid talking with L2 learners (as happens quite frequently, see Derwing, Rossiter, & Munro, 2002), there will be negative consequences, because it is clear that both massive amounts of input and opportunities to speak are necessary to improve L2 fluency. Furthermore, Thomson and Isaacs (2011) found strong correlations of listeners’ judgments of L2 speakers’ intelligence with temporal measures of fluency, suggesting that dysfluent speech can contribute to much broader negative impressions. Thus, improving L2 fluency may contribute to other speakers’ willingness to engage an L2 learner in conversation, providing input and interaction that may lead to increased language learning overall.

Current Issues

The study of L2 fluency is of interest to psychologists, psycholinguists, phoneticians, and applied linguists alike, although each field studies fluency from different perspectives and with different goals. Psychologists and psycholinguists are primarily interested in determining

Tracey M. Derwing

the cognitive processes that affect fluency, while phoneticians generally focus on the physical/acoustic outputs of learners. Applied linguists typically want to identify ways in which learners' fluency can be enhanced through manipulation of tasks in the classroom, the effects of study abroad or other forms of immersion in the L2, the fluency trajectories of learners, and the interrelationships of factors that affect the fluency of utterances produced by L2 speakers. Ultimately, all of this research is linked.

One area that applied linguists have turned to for insights is research on fluency in the L1. In the early 1970s, Pawley and Syder (1975) collected a large body of conversational speech from dozens of native English speakers. As they transcribed the data and examined its characteristics, they formed the hypothesis that speakers operate under a 'one-clause-at-a-time' constraint; that is, they are limited by cognitive processes such as long-term memory to focus on a single clause at a time. Pawley and Syder (2000) revisited their data and expanded their hypotheses, based on dysfluencies such as pausing. They came to the conclusion that "it is the knowledge of conventional expressions, more than anything, that gives speakers the means to escape from the one-clause-at-a-time constraint and that is the key to nativelike fluency" (p. 164). (Pawley and Syder's use of the term 'nativelike' refers to L1 speakers with no language pathologies.) In other words, they noted that there were numerous strings of words that appeared to be learned and used as a single element, and which could thus extend the capacity for fluent production. These strings have been variously referred to as islands of reliability (Dechert, 1980), collocations (Nattinger & de Carrico, 1992), lexical bundles (Biber, Conrad, & Reppen, 1998), formulaic sequences (Wray, 2002), and lexical chunks (Schmidt, 2000). Pawley and Syder (2000) indicated that there are "hundreds of thousands" of "multiword units" (p. 179) available to the average native speaker of English. The now sophisticated development of corpus linguistics allows the identification of such collocations through the use of frequency counts. Several corpora exist (some of which have ongoing contributions) and much of this information is freely available (Davies, 2008). The examination of L1 speech has led to the understanding that speakers take advantage of formulaic sequences to enhance their utterance fluency, regardless of their own processing speed (or cognitive fluency, in Segalowitz's terms). In the case of L2 speakers, there is a comprehensibility benefit of using common collocations, as Wray (2002) has pointed out:

Using word strings which the native speaker, as hearer, can decode easily (because they are formulaic) will greatly enhance the success of the message's interactional purpose, not least because, if the speaker has nonnative-like phonology, the hearer will need to engage in extra processing for the phonological decoding.

p. 99

Furthermore, we now know from a six-month longitudinal study of learners' use of formulaic sequences that they contribute to utterance fluency (measured by mean length of run in syllables) over time (Wood, 2006).

Although learning multiword strings as single chunks is one way of enhancing fluency, clearly, automaticity of other aspects of cognitive processing is another. Several theoretical models have been suggested for application to L2 fluency (see Kormos, 2006, for a comprehensive overview), but one of the earliest is de Bot's (1992) L2 adaptation

of Levelt's (1989) model of the speaking process for unilingual speakers. De Bot argued that, in fact, a model for bilingual speakers should probably be the standard, given that the majority of people worldwide speak more than one language. Both Levelt's and de Bot's models have the same architecture, in that they begin with a 'conceptualizer'—this is the preverbal, semantic notion that the speaker wishes to express. De Bot has argued that for bilingual or multilingual speakers, the choice of language to be spoken is made within the conceptualizer. The next stage in the model is the 'formulator,' where words, grammar, and phonology are implemented. De Bot suggests that there are separate stores within the formulator for each language spoken, but that there is "one common lexicon in which items are connected in networks which enable subsets of items to be activated. One such subset can be the items from a specific language" (p. 14). De Bot introduces a feedback loop here for L2 speakers who often have difficulty with lexical retrieval, so that they may search again for the right word. The 'articulator' is the component in which speech is actually produced, while the information from the formulator is held for monitoring purposes by the 'speech comprehension system,' which can then lead to the correction of any errors that may have occurred. Proficiency in an L2 within this model is a factor that determines fluency; both word-finding speed and knowledge of grammatical and lexical concepts will influence how quickly and smoothly a speaker can produce an intended utterance. Segalowitz (2010) further adapted this model, identifying several points where fluency could break down, including the micro-planning stage in the conceptualizer, the encoding of grammar, lexical retrieval, phonological and phonetic encoding, articulation, and the speakers' self-perceptions of their own productions. The degree of automaticity involved when these functions are put into operation will ultimately affect utterance fluency. As Segalowitz (2013) points out, automaticity can include "speed of processing, stability of processing, the ballistic (unstoppable) nature of the processing and the effortlessness of it" (p. 242). He also suggests that the ease with which the learner can redirect attention as needed will also contribute to overall cognitive fluency. In early stages of language learning, L2 speakers have to actively control their productions, seeking words and structures to express the meanings they intend to relay. If they cannot find a given word, for instance, they may have to resort to paraphrase or some other strategy, a process that will delay their intended utterance, but if their redirection of attention skills is flexible, they may be able to compensate somewhat for their limitations.

The development of L2 fluency has also been approached through the lens of complexity theory, which has been used in recent years to describe L2 development generally (Larsen-Freeman, 2006; Larsen-Freeman & Cameron, 2008). One of the key aspects of complexity theory that differs from other models is that a change (or problem) in one area will have an effect on other areas. Rather than seeing language learning as a linear process, passing from one stage to the next, the linguistic system is dynamic and shifting. Relationships across different components are complex and affect each other. Furthermore, individual trajectories can be quite distinct, because each person's system is affected by context, input, aptitude, and other factors. As Thomson (2015) observes in a discussion of the interrelationships of pronunciation variables and fluency, complexity theory "offers a framework for making sense of the sometimes chaotic evidence for a partial relationship between fluency, accent-ness, intelligibility, and comprehensibility, and opens new directions for fluency and pronunciation research" (p. 221).

Key Concepts

Fluency: Fluency in research contexts generally refers to the flow or fluidity of speech.

Cognitive fluency: The responsiveness and rapidity of the mental processes required to produce speech, such as short-term memory, lexical retrieval, and grammatical choice.

Utterance fluency: Utterance fluency is associated with measures such as speech rate (syllables per second), mean length of run (number of syllables between pauses), number of pauses, mean length of pause, false starts, and repetitions. These features are a result of learners' underlying cognitive fluency.

Perceived fluency: Listener judgments of speaker fluency, typically expressed on a Likert scale ranging from 'very fluent' to 'very dysfluent.'

Formulaic sequences: Also known as 'collocations' or 'chunks,' these are multiword strings that function as a single unit or word.

Automaticity: The extent to which a speaker can engage in several cognitive processing steps without actively allocating controlled attention to each individual step.

Empirical Evidence

As we saw in the discussion on theoretical models, cognitive fluency underlies utterance fluency, which affects listeners' perception of fluency. Cognitive fluency is related to a speaker's aptitude, thus a speaker's L1 fluency on all levels is likely to be reflected in L2 productions as well. Towell, Hawkins, and Bazergui (1996) compared the speech rate (in syllables per minute) of 12 English speakers whose L2 was French. Overall, they found that the faster the speech rate was in the L1, the faster the rate in the L2. Derwing, Munro, Thomson, and Rossiter (2009) compared L1 and L2 fluency ratings for 16 Mandarin and 16 Slavic language speakers in a two-year longitudinal study. The speakers produced narratives from a set of cartoons depicting a story about two strangers who accidentally took each other's suitcase home. At the outset of the study, the speakers were enrolled in ESL classes; a standard test of speaking and listening indicated that they were all at the same level of oral proficiency. They were asked to produce the suitcase narrative in their L1s to obtain baseline data, and subsequently, narratives were collected in English at the 2-month, 10-month, and 2-year points. Segments of the L1 narratives were played to eight native Mandarin listeners and eight native Russian listeners. The L1 listeners were told that they would use a 9-point scale, ranging from extremely fluent to extremely dysfluent, to assess the voices, and that they should use as much of the scale as possible, but that there was no expectation for them to use the whole scale because the speakers were normal in the sense that they did not exhibit any pathologies such as a stutter or dysarthria. (None of the speakers was extremely dysfluent in the L1, so it would be inappropriate for raters to use the whole scale, despite noticeable individual differences in fluency across speakers.) The listeners were told that the length and number of pauses, self-repetitions, and false starts were how they were to gauge L1 fluency. Another eight raters, native speakers of English, listened to the L2 samples, which had been randomized for time, and rated them for fluency. These raters too had been told to focus on temporal factors, rather than proficiency. When correlations of ratings for L1 and L2 fluency from the two-month data collection period were conducted, significant relationships were found for both the Mandarin and Slavic language groups, as expected. However, when comparisons of the L1 and L2 fluency ratings at two later collection times were made, there were

no significant relationships. The authors concluded that a range of factors may have influenced the findings, including greater variability in English proficiency across participants; however, they argued that although complex, there is a relationship between L1 and L2 fluency, such that one would not expect a slow talker in L1 to exhibit greater fluency in the L2.

In a study comparing lower and higher proficiency Korean speakers of English as an L2, Kahng (2014) examined cognitive fluency by first eliciting 2 minutes of spontaneous speech, and then conducting spontaneous recalls in which the participants commented on difficulties in their productions. Kahng reasoned that lower proficiency learners would likely remember more than higher proficiency individuals, partly because they would be more dependent on declarative (consciously stored) knowledge; that is, they would exhibit less reliance on automated or procedural knowledge. Indeed, not only did the lower proficiency speakers remark more frequently on their production difficulties, but their comments were qualitatively different, in that they focused more on specific vocabulary and syntactic issues, referencing lexical retrieval problems and specific grammar rules.

The L1 is reflected in L2 oral fluency in another sense, given that in Derwing et al. (2009), and in several other studies, Derwing and her colleagues found that Mandarin learners of English in an immigrant environment made consistently fewer gains over time in fluency than did Slavic language speakers (primarily Russians). Even after seven years in a largely English-speaking environment, the Mandarin speakers showed little or no improvement on fluency rating tasks, whereas the Slavic language speakers were perceived as being significantly more fluent from the 2-year point to the 7-year point (Derwing & Munro, 2013). A closer examination of individual trajectories between the 2-year point in the study and 7 years later indicated that only 4 of the 11 Mandarin speakers showed any perceived fluency improvement in that time period, whereas 8 of the 11 Slavic language speakers were judged to be more fluent. It is conceivable that L1 was indeed a factor here, in that English and the Slavic languages (Russian and Ukrainian) are both Indo-European whereas Mandarin is unrelated to English.

On the other hand, L1 may also serve as a proxy for other factors that influenced the differences in oral fluency. In interviews with many of the same individuals who participated in the Derwing and Munro (2013) study previously cited, Derwing, Munro, and Thomson (2008) asked them to indicate on a 5-point scale ranging from ‘never’ to ‘several times a day’ how often they had conversations of 10 minutes or more in English. They determined that the Mandarin speakers spent significantly less time interacting in their L2. Similarly, the participants were queried about the time spent listening to talk radio in English. Again, the Slavic language speakers spent significantly more time deliberately exposing themselves to the L2 as indicated in this comment “I’m listen to radio every morning when I’m driving to school. I force myself. . . . I totally deprive myself of Russian movies . . . I started enjoy watching English movies, you know . . . even I go to the theatre, now I can understand at least 70%” (p. 372). The authors used the Willingness to Communicate (WTC) framework developed by MacIntyre, Clément, Dörnyei, and Noels (1998) to interpret the results. As MacIntyre et al. (1998) indicated, “certain groups may be more homogeneous than others with respect to certain traits or profiles. As well, groups may show different average or baseline levels of a given trait” (p. 558). Using interview data, Derwing et al. (2008) determined that the differing L2 fluency levels of the two language groups were well explained by socio-affective factors, and motivational matters. There were similarities between groups, for instance, such as the greater willingness to initiate a conversation on the part of Slavic language speakers, despite a fear of making mistakes, and closer connections to the local Chinese community on the part of the Mandarin speakers, meaning less time spent with English speakers.

Tracey M. Derwing

Another factor that affects utterance fluency is the nature of the linguistic task. Derwing et al. (2004) compared three tasks, a picture narrative, a monologue, and a conversation. In listener judgments, the picture narrative was considered to be significantly less fluent than the other conditions, a finding that the authors attributed to the nature of the picture description task, which allowed less lexical choice than either of the other tasks. In both the monologue and the conversation, speakers could avoid lexical items and structures that were difficult, and could focus on relatively familiar information. Foster and Skehan (1996) also found differences in fluency related to task. Their participants were less fluent in a picture description task than a personal information exchange or a collaborative decision-making activity. The authors argued that the source of the difference in fluency across tasks was a difference in cognitive load. The picture narrative forced the speakers in certain directions in which they were unable to capitalize on the freedom and familiarity that the other tasks offered.

The opportunity to plan before undertaking a speaking task has also been found to have an impact on oral fluency (Ellis, 2009). In a review of several studies of L2 pretask planning, Ellis found that rehearsal had a positive effect on fluency, but it did not necessarily transfer to a new task in the absence of additional interventions. Ellis identified 19 studies that examined strategic planning, in which learners were given time to prepare for an oral task. In 17 of the 19 studies, planning contributed to increased fluency. The studies differed in many ways, making direct comparisons difficult, but a few generalizations can be made. Length of planning time matters such that longer is better; 5–10 minutes is preferable to a single minute, for instance. However, some learners did not avail themselves of all the planning time allotted to them because they did not see the value. And finally, more research should be conducted to gauge the value of guided versus unguided planning. Ellis suggests that the benefits of these factors may depend on the nature of the task.

The dynamic between an L2 speaker and an interlocutor is another factor that can have a powerful effect on fluency. In a study of the instruction of pragmatics, Derwing, Waugh, and Munro (2014) tested the hypothesis that improved control of culturally determined pragmatic language would result not only in stronger ratings of cultural appropriateness, but also in improved perceptions of fluency on the part of listeners. The researchers asked intermediate ESL speakers to participate in audio-recorded role play scenarios both before and after instruction. The scenarios were designed to elicit four speech acts: requests, refusals, compliments, and apologies. Over the course of 5 weeks, the learners received 25 hours of instruction, including analysis of videos portraying inappropriate linguistic behavior, and analyses of the students' own role plays. In addition, they undertook standard pragmatic activities such as scenario completion in small groups, listening to notice certain forms such as softeners (e.g., *I just need a minute*) and intensifiers (e.g., *I'm really sorry*), and overt instruction of formulaic sequences. Pre- and postaudio samples of two refusal and two request scenarios were randomized and played to 56 listeners, who rated the speakers on 9-point scales for fluency and social appropriateness. In all four scenarios, the listeners judged the postinstruction condition to be significantly more socially appropriate. In other words, the pragmatics course was effective. However, contrary to expectation, there was no significant improvement in fluency on two of the scenarios, and a significant decrease in fluency in another. Significant improvement in the perception of fluency was found in only one case. Interestingly, in the three situations where there was no improvement or a worsening in fluency, the L2 speakers were playing roles where they were not in power. In the worst case setting, they were to remind their boss that he/she had promised a raise after 3 months on the job. It was now the fourth month, and the L2 speakers had to ask for a raise. In the other

two situations, where no difference in fluency appeared over time, the L2 speakers had to ask their employer if they could leave early to pick up their children from school, and they had to refuse their employer's request to work overtime. In the final situation, where significant improvement in fluency was perceived, the L2 speakers took the role of bank clerk, and refused to serve a customer who had no ID. This study suggests not only that the power dynamics between interlocutors can affect fluency, but that the complexity of the message to be relayed influences it as well.

Finally, overt instruction aimed at improving fluency is an obvious factor that will be considered in the next section.

Pedagogical Implications

Although speaking fluency is limited by the learner's own cognitive processing speeds, regardless of the learner's abilities, pedagogical activities can increase automaticity through increased awareness of fluency markers, planning and rehearsal tasks, the instruction of frequently occurring formulaic sequences, common discourse markers, and an intensified focus on general speaking and listening tasks.

There is a widespread consensus that many L2 students do not have much opportunity to enhance their spoken fluency in classrooms. Several factors militate against fluency practice, including large class sizes, competing demands of other language skills that must be taught, time limitations, and, in some instances, a lack of familiarity on the part of the teacher with activities that target oral fluency. Studies suggest that to become more fluent in production, learners need to practice speaking; they are likely to improve if they engage in substantive interactions outside the classroom (Derwing et al., 2008). However, if they are not motivated to do so on their own, or if they do not know how to insert themselves into situations that allow them to speak with others beyond familiar routines, students would benefit both from more focus on oral fluency development in the classroom and help accessing interaction opportunities in the L2 community.

In terms of instructed oral L2 fluency, there are very few studies. Perhaps one of the earliest and best known is that of Nation (1989), who showed that repeated tellings of the same story in progressively shorter periods of time resulted in fewer pauses and fillers; that is, the learners became more fluent performing the same task under the pressure of reduced time. He used a technique outlined in the Teaching Tips section of this chapter, and measured students' fluency at each repetition of the task, showing that learners were more fluent with each retelling.

Temple (2005) investigated fluency development in instructed learners of French. She collected data from pairs of learners interviewing each other both before and after 3 months of instruction. Samples of speech from the interviews were analyzed for pause phenomena, speech rate, and hesitations, which included incomplete words, repeats, and false starts. Eight of the 11 participants showed improvement in speech rate postinstruction. They had not simply sped up their productions but their pause placements also differed, such that the early interviews exhibited many clause-internal pauses, while the more fluent interviews postinstruction had more clause-initial pauses, indicative of planning. The author concludes that pause placement at the outset of a clause is more native-like, and indicative of an increase in cognitive fluency. That is, the gains in fluency demonstrated more automatic versus controlled behavior. This study is reminiscent of Pawley and Syder's (2000) one-clause-at-a-time hypothesis, in that the learners were progressing from less than a clause to a full clause in formulating their productions.

Tracey M. Derwing

Gatbonton and Segalowitz (2005) proposed a framework for teaching fluency called ACCESS—or “Automatization in Communicative Contexts of Essential Speech Segments” (p. 328), which has three phases: creative automatization, language consolidation, and free communication. Using the example of the theme *Family*, the authors describe each of the phases. The first consists of a pretask and the main task. The pretask is designed to determine whether the students have enough vocabulary/phrases to attempt the main task. The main task must be “genuinely communicative, inherently repetitive, and functionally formulaic” (p. 331). Two groups are formed, and are told that they must develop a family and the history that that entails. Students work together, eliciting information from each other necessary to complete the overall goal of the task. In their example, the authors propose that students decide their own roles in the family they are creating. Students from one group interview those from the other in order to present the family organization to the whole class. Essential speech elements (the target utterances or formulaic sequences chosen by the instructor) are introduced to suit the task, but these elements can also be used readily outside the classroom. In the consolidation phase, the teacher reinforces important sequences, and conducts form-focused instruction as necessary. The free communication phase allows the learners to discuss their chosen topic, but because they have been working with target phrases all along, they are likely to repeat many of the essential speech elements often, thus enhancing fluency.

Nation and Newton (2009) provide suggestions based on research for a listening and speaking pedagogical approach to enhance L2 fluency. They argue that to become fluent, learners must focus on all four language skills. Tasks centered on fluency must entail language with which the learners are completely familiar. The focus should be on receiving and sending messages with some pressure to speak quickly. Furthermore, Nation and Newton argue that learners need large amounts of input and should be encouraged to produce similar amounts of output to become more fluent. They adhere to Swain’s (1985) Output Hypothesis, that it is through producing utterances in an L2 that the learner can “move from a purely semantic analysis of the language to a syntactic analysis of it” (Swain, 1985, p. 252). In other words, having to speak requires students to notice aspects of the L2 that are not necessary for comprehension. Nation and Newton propose that learners be required to practice formal speaking, which entails longer turns, demanding greater fluency. They offer many specific suggestions for classroom tasks, some of which appear in the Teaching Tips section of this chapter.

Another resource intended for L2 teachers who want to address the fluency of their learners is Rossiter, Derwing, Manimtim, and Thomson’s (2010) article, which outlines a full range of classroom activities that can be adapted to different proficiency levels. Many of the Teaching Tips in this chapter are described in detail in Rossiter et al.’s article; most of the activities outlined here are based on evidence from research on the development of L2 fluency.

A recent study comparing two forms of instruction for their effect on fluency development is possibly the clearest indication yet that teaching tailored to oral production can enhance L2 fluency in learners. Galante and Thomson (2016) examined the performance of Brazilian pre-intermediate learners of English as a foreign language (EFL) on five tasks both before and after a 4-month instruction period (74 hours of teaching). Half the students were taught in traditional, communicative language classrooms, and were given pair and group work typical of communicative teaching. Furthermore, these students were required to research, prepare, and deliver oral presentations to the rest of the class at the end of the semester on a topic of their own choosing. The other students were taught using drama

methods, including problem solving, role plays, and a short play performed at the end of the term. The drama activities encouraged considerable improvisation rather than the use of the scripted material so frequently found in communicative textbooks (cf. Diepenbroek & Derwing, 2014). Pre- and posttreatment speech samples from the five oral tasks were randomized and played to 30 Canadian university students, none of whom had familiarity with a Portuguese accent. The raters assessed the fluency of each speech sample on a 9-point scale (1 = very fluent; 9 = very dysfluent). The comparisons of pretest and posttest for each group indicated that there were no significant differences in the judgments of the two groups of learners prior to instruction, but that the students who had been taught using drama methods were perceived to be significantly more fluent on the posttests, whereas there was no change in the perceived fluency of the students in the communicative classroom at the end of the research period. The authors point out that although studies such as that of Nation (1989) show that repetition can result in improved fluency within a given activity, the use of drama techniques led to increased fluency in five distinct oral production tasks. The students had evidently been able to generalize their improvisational skills.

Teaching Tips

- Raise students' awareness of markers of fluency, such as appropriate intonation to indicate that the speaker still holds the floor; placement of pauses at phrase or clause boundaries; and explicit instruction of oral fluency. One way to do this is to have students transcribe short YouTube videos, which the class can then analyze together, and then practice with shadowing and role play activities. Shadowing is a popular technique for helping students recognize markers of fluency. Students should have a transcript of a short video- or audio-recording, in order to read along with the voice to be modeled. They can either speak at the same time as the recording, or just slightly after.
- Guillot (1999) recommends having students watch talk shows in their L2, to analyze the speakers' use of strategies for buying time, without sounding dysfluent.
- Engage students in rehearsal and repetition tasks, such as Nation's classic (1989) 4–3–2 task (sometimes called fluency circles), in which learners recount a story to a classmate in 4 minutes, then tell the same story to another classmate in 3 minutes, and finally, repeat the same account to a third student in 2 minutes.
- Develop activities that focus on meaning-making.
- Explicitly teach high frequency formulaic sequences appropriate to students' proficiency level (see <http://corpus.byu.edu/> for a listing of formulaic sequences by frequency).
- Teach discourse markers such as fillers (e.g., 'like,' 'you know,' 'so,' and 'well') sequential markers (e.g., 'first,' 'next,' 'then,' and 'finally'), and conventions for opening and closing a conversation.
- Provide students with contact activities in which they must interact with others in their L2 (e.g., conducting a short, in-person survey with at least five speakers who are not classmates; phoning a call center for information).
- To practice formal speaking, Nation and Newton (2009) suggest the 'pyramid procedure.' Have a student prepare notes for a talk, then give the talk to a fellow student for feedback. After revising and shortening the notes, the student should present again to a small group of students for additional feedback. Finally, have the student present to the whole class, using very few notes.

Tracey M. Derwing

One approach to accessing the L2 in authentic settings for gains in fluency is Study Abroad programs. Lennon (1990) conducted one of the earliest study abroad research projects, in which he compared fluency ratings with temporal measures for four German learners of English who spent six months at a university in England. The participants produced picture narratives at the beginning and end of their stay, which were subsequently transcribed and analyzed for 12 different measures of utterance fluency, including words per minute, repetitions, self-corrections, percentage of unfilled pause time, percentage of filled pause time, and mean length of run. All four participants showed improvement on at least some of the fluency measures; in particular there was a reduction in pause time. Furthermore, Lennon noted, as other researchers would find in later studies (e.g., Wennerstrom, 2001), that pause placement was important, as well as length and frequency of pauses. Lennon also had 10 judges assign global fluency ratings to the pre- and postproductions. Overall, they agreed that each learner became more fluent, although there were some disagreements among judges in the case of each participant.

In a special issue of *Studies in Second Language Acquisition* (2004, issue 2) edited by Joseph Collentine and Barbara Freed, several researchers explored the development of fluency in Study Abroad settings. Their findings indicated that simply going abroad to a country where the L2 is spoken does not guarantee that a learner will make fluency gains. Much depends on the context in which the learner is embedded, the learner's own willingness to communicate, and the proficiency level of the learners on arrival. Segalowitz and Freed (2004), for example, examined two comparable groups of learners of Spanish; one group studied at home in the USA, while the other group spent a semester in Spain. Both sets of learners were asked about language contact, and the study abroad group was also questioned about the amount of time they spent with their home-stay family. Along with measures of cognitive processing, the students' utterance fluency was measured pre- and poststudy abroad by examining speech rate, number of hesitations, number of fillers, and the "number of words in the longest fluent speech run" (p. 183). Overall, the study abroad group improved significantly on three of the four measures, while the at-home group showed no significant changes in oral fluency over time. However, the authors conducted additional analyses in which the extracurricular activities reported by the study abroad students were co-varied out, thus indicating that the opportunities to have more contact in the L2 did not necessarily contribute to their better fluency performance. Interview data suggested that, in fact, the more contact with host families the students had, the worse their oral performance. The authors surmise that the nature of the conversations within the host family context may have been restricted to short, somewhat banal interactions, thus limiting the learners' productions. In other words, the interactions were the same, routine, politeness conventions that did not extend to true conversations. The authors further propose that there may have been a threshold of proficiency and cognitive skills that would allow them to take advantage of opportunities to speak. The findings are reminiscent of the Willingness to Communicate framework described in MacIntyre (2007) and MacIntyre et al. (1998), which outlines the complex array of factors that contribute to a learner's initiation of an interaction.

In ESL contexts, practice outside the classroom in authentic settings is another route to improving L2 fluency. One approach that has merit is the use of volunteer opportunities that benefit both the organization involved and the L2 student. Dudley (2007) surveyed 55 adult students in an ESL program to determine whether they had volunteered in the community, and, if not, whether they would consider volunteering in the future. Forty-six of the students indicated that they had not volunteered, primarily because of "a lack of opportunity and knowledge about volunteering" (p. 546), but of those individuals, 87% indicated that they would like to volunteer in the future. Dudley also interviewed those learners who had found volunteer positions; most had opportunities to interact in their new language and were

able to gain work experience in their new culture, but a few were placed with others who spoke their L1, and some had little chance to speak with others in their placements. Dudley recommended that ESL programs serve as a liaison between organizations that could offer volunteer experiences and L2 learners, ensuring both that learners are not exploited, and that they gain interaction opportunities commensurate with their needs.

In addition to identifying potentially interaction-rich volunteering opportunities for learners, as Dudley suggested, Derwing and Waugh (2012) recommended a different venue for authentic interaction opportunities. For many years, the federal government of Canada has provided funding to settlement agencies to run what was originally called a *Host Program*, and what is now known as *Community Connections*. Agencies enlist high proficiency or native speakers of English as volunteers who spend time with newcomers on a social basis. These programs help L2 learners gain practice speaking with others in English in genuine conversations. Other immigrant-receiving countries with large numbers of language learners would do well to consider similar programs. In fact, Yates et al. (2010), in a policy report to the Australian government, recommended not only that language classes should direct “explicit attention to language learning and social networking strategies” (p. 80), but that the government should develop and promote “community outreach programs to increase awareness in the broader community of migrant issues and strategies for interacting with speakers from different language backgrounds, in particular programs that bring expert speakers of English and newly arrived migrants together” (p. 80).

Future Directions

Research has pointed to several ways in which learners’ utterance fluency can be enhanced, but typically, we focus on only one type of fluency at a time, usually utterance fluency. Segalowitz (2010) has recommended that researchers consider fluency in a holistic manner, taking all three types of fluency into account. This suggests a programmatic approach, in which teams of researchers collaborate in planning and executing connected studies. Psycholinguists, psychologists, applied linguists, and L2 teachers all have a role to play here. In the meantime, we also need more research on the outcomes of pedagogical interventions over the long term. As is the case with most applied linguistics research, there is a need for more longitudinal research that traces fluency development over extended periods of time to examine individual trajectories. We now also have the technology with mobile phones to record natural conversations as Surtees (2015) did. Such recordings would give researchers and instructors a window on how often students initiate a conversation, and whether they employ means for extending exchanges in real life settings. This information could then feed into suggestions for classroom work. Although some studies suggest that fluency can be enhanced by particular classroom tasks, and that naturalistic practice results in increased fluency, detailed comparisons of various activities could also help to determine which are most effective, and an understanding of individual propensities could lead to instruction that is tailored to learners’ own needs.

Acknowledgments

I thank the editors for their very helpful comments. I am most grateful to Ron Thomson, who provided useful input on an earlier version of this chapter. Finally, I thank Murray Munro for his ongoing contributions to my conceptions of language teaching. Any errors are, of course, my own.

Tracey M. Derwing

References

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. New York: Cambridge University Press.
- Davies, M. (2008). The corpus of contemporary American English: 520 million words, 1990-present. Retrieved from <http://corpus.byu.edu/coca/>
- de Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics*, 13, 1–24.
- Dechert, H. W. (1980). Pauses and intonation as indicators of verbal planning in second language speech productions: Two examples from a case study. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 271–285). The Hague: Mouton.
- Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, 63, 163–195.
- Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29, 359–380.
- Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533–557.
- Derwing, T. M., Rossiter, M. J., & Munro, M. J. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multicultural and Multilingual Development*, 23, 245–259.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54, 655–679.
- Derwing, T. M., & Waugh, E. (2012). Language skills and the social integration of Canada's adult immigrants. In *IRPP study #31*. Montreal: Institute for Research on Public Policy.
- Derwing, T. M., Waugh, E., & Munro, M. J. (2014, March). *Willingness to communicate and L2 speakers' pragmatic development: Implications for instruction*. Paper presented at the American Association for Applied Linguistics, Portland, Oregon.
- Diepenbroek, L. G., & Derwing, T. M. (2014). To what extent do popular ESL textbooks incorporate oral fluency and pragmatic development? *TESOL Canada Journal*, 30, 1–20.
- Dudley, L. (2007). Integrating volunteering into the adult immigrant second language experience. *Canadian Modern Language Review*, 63, 539–561.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 30, 474–509.
- Fillmore, C. J. (2000). On fluency. In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 43–60). Ann Arbor, MI: University of Michigan Press.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299–323.
- Galante, A., & Thomson, R. I. (2016). The effectiveness of drama as an instructional approach for the development of L2 fluency, comprehensibility, and accentedness. *TESOL Quarterly*. doi:10.1002/tesq.290
- Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, 61, 325–353.
- Guillot, M.-N. (1999). *Fluency and its teaching*. Clevedon: Multilingual Matters.
- Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, 64, 809–854.
- Kormos, J. (2006). Fluency and automaticity in L2 speech production. In J. Kormos (Ed.), *Speech production and second language acquisition* (pp. 154–165). Mahwah, NJ: Lawrence Erlbaum.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27, 590–619.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387–417.

- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *Modern Language Journal*, 91, 564–576.
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. (1998). Conceptualizing willingness to communicate in an L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, 82, 545–562.
- Nation, P. (1989). Improving speaking fluency. *System*, 17, 377–384.
- Nation, P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.
- Nattinger, J. R., & de Carrico, J. S. (1992). *Lexical phrases and language teaching*. London: Longman.
- Pawley, A., & Syder, F. H. (1975). Sentence formulation in spontaneous speech. *New Zealand Speech Therapists' Journal*, 30(2), 2–11.
- Pawley, A., & Syder, F. H. (2000). The one-clause-at-a-time hypothesis. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 163–197). Ann Arbor, MI: University of Michigan Press.
- Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review*, 65, 395–412.
- Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *Canadian Modern Language Review*, 66, 583–606.
- Schmidt, N. (2000). Lexical chunks. *ELT Journal*, 54, 400–401.
- Segalowitz, N. (2010). *The cognitive basis of second language fluency*. New York: Routledge.
- Segalowitz, N. (2013). Fluency. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 240–244). New York: Routledge.
- Segalowitz, N., & Freed, B. (2004). Context, contact and cognition in oral fluency acquisition. *Studies in Second Language Acquisition*, 26, 173–199.
- Surtees, V. (2015, March). *Going beyond the classroom: Participants' mobile devices as research and learning tools*. Paper presented at American Association of Applied Linguistics, Toronto, Ontario.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–256). New York: Newbury House.
- Temple, L. (2005). Instructed learners' fluency and implicit/explicit language processes. In M. Pierrard & A. Housen (Eds.), *Investigations in instructed second language acquisition* (pp. 31–49). Berlin: De Gruyter Mouton.
- Thomson, R. I. (2015). Fluency. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation* (pp. 209–226). Malden, MA: John Wiley & Sons.
- Thomson, R. I., & Isaacs, T. (2011, March). *Speaking in a drone: Are listeners attuned to pitch-related differences when rating L2 speech samples for personality attributes?* Paper presented at American Association of Applied Linguistics Conference, Chicago, IL.
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84–119.
- Varonis, E., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, 4, 114–136.
- Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. New York: Oxford University Press.
- Wood, D. C. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review*, 63, 13–33.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Yates, L., Ficorilli, L., Kim, S. H. O., Lising, L., McPherson, P., Taylor-Leach K., Setijadi-Dunn, C., Terraschke, A., & Williams, A. (2010). *Language training and settlement success: Are they related?* Research report to the Department of Immigration and Citizenship, Australia. Sydney: Adult Migrant English Program Research Centre.

Pronunciation Acquisition

Sara Kennedy and Pavel Trofimovich

Background

It is only in the last few decades that research on instructional interventions for second language (L2) pronunciation has developed into a recognizable research area. The historical scarcity of research on L2 pronunciation instruction has been explained in several ways. Murphy and Baker (2015) cite the strong focus on L2 reading and writing instruction in the 1970s and 1980s, a period when research on L2 development was proliferating. Derwing and Munro (2015) attribute the shortage of research to a combination of factors: the influence of Purcell and Suter's (1980) cross-sectional study showing no effect of pronunciation instruction on L2 accent, the widespread adoption of Krashen's ideas (e.g., Krashen, 1989) about the importance of comprehensible input over explicit instruction in the classroom, and the growing pedagogical emphasis in the 1980s on the value of activities that foregrounded communication over linguistic accuracy.

Whatever the reasons for the paucity of research, even a century ago pronunciation specialists were concerned with issues that are important to researchers today. These issues include considering speakers' intelligibility as a norm for acceptable pronunciation (Sweet, 1900), the relationship between speakers' identity and their pronunciation (Abercrombie, 1949), and learner autonomy and self-monitoring (Sisson, 1970). However, until the 1990s, most published work in L2 pronunciation teaching was not research-based but consisted of position papers, methodology guides, and teaching and learning materials, which were based on theoretical reasoning, anecdotal evidence, or individual teachers' experiences (Murphy & Baker, 2015). At present, many pronunciation teaching issues are being explored through research, although several areas, such as learner identity, early L2 classroom learners, classroom corpora, and teacher education remain under- or uninvestigated. In this chapter, we critically examine selected areas of research relevant to L2 pronunciation instruction.

Pronunciation and Other L2 Skills

L2 pronunciation learning is different from the learning of other L2 skills such as writing (Fraser, 1999). First, practically all L2 users who were adults before their first or most

extensive exposure to the spoken L2 show some evidence of nonnative pronunciation in spontaneous production, even after intensive instruction (Derwing & Munro, 2013). Yet at least some L2 users who were immersed in an L2 environment from childhood (e.g., as immigrants) speak with nativelike accents, often without pronunciation instruction (Abrahamsson & Hyltenstam, 2009). Second, L2 speech research shows a consistent link between speakers' perception and production, such that learners who struggle to accurately perceive L2 segments or prosody also struggle to accurately produce them (Flege, 2003). These two observations have shaped research on L2 pronunciation instruction, in terms of the age of participants targeted and the instructional and methodological focus on both perception and production.

Theoretical Frameworks

Until the beginning of the 21st century, almost no research on L2 pronunciation instruction drew on theoretical frameworks for L2 learning or teaching (Munro & Derwing, 2011). The end of the 19th century saw a growing number of theoretical descriptions of phonological systems, and in the 1950s, linguistically based hypotheses about pronunciation learning challenges (e.g., Contrastive Analysis) as well as teaching approaches based on theoretical frameworks (e.g., Audio-Lingual Method based on Behaviorism, Silent Way based on Cognitive Code) began to appear. Although the relevance of these views for pronunciation learning was not investigated in research, they were nevertheless adapted for and incorporated into various teaching and learning materials, influencing the selection and description of instructional targets (Derwing & Munro, 2015; Murphy & Baker, 2015).

Beginning in the 1980s, the learning of L2 pronunciation (albeit with a nearly exclusive emphasis on individual segments) was the focus of extensive research that resulted in several complementary theories, including the Speech Learning Model (Flege, 2003) and the Perceptual Assimilation Model (Best & Tyler, 2007). Although many researchers have invoked these theories to provide a conceptual backdrop for their research and to discuss its findings, very few have directly examined the predictions of these and other theoretical frameworks for L2 pronunciation instruction. This deficit may be partially due to the often loose links between L2 speech research and pronunciation instruction, with theoretical views often having little to contribute to pedagogy (Fraser, 2004), either because the views were not designed with practice in mind or because such links have not yet been established.

Pedagogical Norms in Research

As Munro and Derwing (2008) note, in the majority of research targeting L2 speakers' pronunciation, development has been assessed through acoustic measures, such as formant frequency or duration measurements, or through listener-based ratings of accent. The standard of reference has been native speakers' pronunciation (e.g., a range of nativelike values for acoustic measures or listeners' reference to the pronunciation of imagined native speakers). Early studies on L2 pronunciation instruction typically adopted the same approach, with pronunciation development measured through judgements of natelikeness or accentedness (e.g., de Bot & Mailfert, 1982). Implicit in these measures is the idea that the default norm for L2 pronunciation is a native speaker norm and that the aim of L2 pronunciation instruction is to lessen or eliminate any traces of nonnativeness.

However, in the early 1980s, researchers began to explore other means of assessing L2 pronunciation, such as measuring listener understanding of L2 speakers (e.g., Varonis & Gass, 1982). Two influential studies in this vein were Derwing and Munro (1997) and Munro and Derwing (1995), which introduced three interrelated constructs: intelligibility, comprehensibility, and accentedness. Intelligibility was defined as “the extent to which a speaker’s message is actually understood by a listener” (Munro & Derwing, 1995, p. 76), measured in this instance through listeners’ transcription of L2 speech, while comprehensibility referred to listeners’ “judgments on a rating scale of how difficult or easy an utterance is to understand” (Derwing & Munro, 1997, p. 2). Accentedness was not defined in Munro and Derwing (1995) or in Derwing and Munro (1997), but was measured through listeners’ scalar ratings. These studies revealed partial independence between these constructs, especially between accent and intelligibility, such that listeners can rate some speakers as strongly accented but still clearly understand them. These findings brought into question the use of nativelike pronunciation, whether acoustically or perceptually defined, as the sole norm for assessing L2 pronunciation and stimulated a debate about the appropriate targets and effective methods for pronunciation instruction (Gilbert, 2001; González-Bueno, 1997).

This debate received new energy from Jenkins (2000) and other researchers who work on the pronunciation of English used as a lingua franca, following a prominent tradition of sociolinguistic research (discussed later), which highlights extensive variability among native and nonnative users of English (Kachru, 1992). Many scholars in the field of English as a lingua franca have argued that communication in English now most frequently occurs between speakers who do not claim English as a first language (L1). Therefore, pronunciation instruction and assessment should focus on those aspects of pronunciation that cause problems for understanding between L2 English speakers, and should not center on native speakers’ understanding (Seidlhofer, 2011). This argument reflects the notion of multicompetent speakers, who should not be considered “deficient monolinguals” (Cook, 1992, p. 577) but speakers who succeed in using other languages in addition to their previously learned language(s). To date, L2 pronunciation researchers have adopted a variety of norms for pronunciation measures, from those focusing on native speaker pronunciation to those highlighting native and L2 listeners’ understanding.

Key Concepts

Pronunciation: various dimensions of spoken language, encompassing segments (individual vowels and consonants) and prosody (e.g., word stress, intonation).

Nonnative: demonstrating the influence of a language other than the target language, often in relation to a range of values for acoustic measures or listeners’ rating based on a sample of imagined reference speakers.

Accentedness: degree to which a speaker’s L2 accent resembles that of a given speaker community (often native speakers of the target language).

Approach: set of beliefs and principles used to teach a language (e.g., Behaviorism).

Method: instructional design based on a particular approach (e.g., Audio-lingualism).

Pedagogical norm: language forms that serve as targets for learners to acquire.

Pronunciation model: set of pronunciation forms for a given language variety (accent), often used as reference for instruction.

Current Issues

What Are Appropriate Pedagogical Norms and Targets?

The pedagogical norms, and associated instructional targets, appropriate for L2 pronunciation instruction are still under debate. Levis (2005) categorized approaches to teaching and assessing L2 pronunciation according to one of two principles: the nativeness principle and the intelligibility principle. In approaches drawing on the nativeness principle, the learning and use of nativelike, unaccented pronunciation is the primary aim of instruction. In approaches drawing on the intelligibility principle, the primary aim is for learners to be understood by listeners; nativelike pronunciation by adult learners is viewed as unnecessary and impossible for all but a few. As Derwing and Munro (2015) note, measures of accent cannot stand in for measures of listener understanding, which is a fundamental requirement for successful communication.

Even for those who adopt the nativeness principle, there is still a question of which native variety to use as a pronunciation model. In terms of English varieties, Kachru (1992) famously described the diffusion and use of English around the world as a series of concentric circles. Inner Circle countries (e.g., England, the US) are those where English has traditionally been the primary language; varieties from these countries are often preferred as models in L2 English instruction, including for pronunciation (see Kang, 2015 for overview). In Outer Circle countries (e.g., Singapore), English learning and use has spread over the last few centuries, typically through colonization; from childhood, multiple generations of English speakers have used local varieties of English in personal and public settings. Selecting a particular native variety as a pronunciation model is sometimes a conscious choice by teachers or learners. However, as Derwing and Munro (2015) note, the variety that is most frequently modeled for learners in language classrooms, intentionally or not, is usually that of their teachers.

With regard to the intelligibility and nativeness principles, we maintain that in some very particular contexts of language teaching, learning, and use, the adoption of nativelike pronunciation as a pedagogical norm may be justifiable; however, it is crucial to determine how listeners will understand learners' pronunciation, so research on L2 pronunciation instruction must include listener-centered measures of understanding.

Who Is the (Imagined) Interlocutor?

In L2 pronunciation instruction, a basic but often implicit assumption is that learners will be speaking to listeners or interlocutors, whose reactions to and understanding of L2 pronunciation are an important gauge of learner performance. In research on L2 English pronunciation, the linguistic status of assumed or imagined interlocutors has become the subject of spirited discussion. In her book on English as a lingua franca, Jenkins (2000) argued that the majority of English users today are L2 speakers; therefore, using native English standards in instruction and assessment of L2 English pronunciation makes little sense. Jenkins put forward what she called the Lingua Franca Core, a revised syllabus for English pronunciation instruction based on data she had collected on communication breakdowns and accommodation between nonnative English interlocutors from different language backgrounds. For Jenkins, one implication of this syllabus was that teachers should focus only on those items that were important for intelligibility between L2 users.

Even if learners' pronunciation was nonnative in other aspects, teachers should view those aspects as acceptable as long as intelligibility is achieved. This proposal has generated much debate among researchers and professionals in L2 English pronunciation instruction, with many suggesting that it runs counter to many learners' goals of developing pronunciation that is more than just intelligible to listeners (e.g., Wach, 2011). Research on pronunciation in other L2s (besides English) has rarely included L2 listeners in pronunciation measures; however, in several recent studies on L2 pronunciation development in French and German, L2 listeners have participated. O'Brien (2014) found that L2 German listeners' ratings of L2 German speakers could be predicted by many of the same linguistic measures that were important for native listeners, but unlike past findings for native listeners, ratings of accentedness and fluency (smoothness of speech) were also predicted by measures of grammatical and lexical accuracy. Kennedy, Guénette, Murphy, and Allard (2015) showed that for pairs of L2 French speakers engaged in interactive tasks, pronunciation-related difficulties in understanding were linked primarily to speakers' production of individual sounds.

We consider the issue of pedagogical norms (e.g., nativelike vs. intelligible speech) to be separate from the issue of imagined interlocutors (e.g., measuring L2 speech through judgments by native listeners or by L2 listeners). However, the common practice in L2 pronunciation research of using only native listeners as raters or interlocutors seems shortsighted (see Crowther, Trofimovich, & Isaacs, 2016, for a different approach). Especially for target languages that are major global or regional languages (e.g., English, Spanish, Mandarin), it is irresponsible to presume that in the future L2 learners will exclusively speak with native speakers of those languages.

Which Linguistic Dimensions of L2 Speech Are Relevant to Listener Understanding?

If listener understanding is important for successful communication, then identifying linguistic barriers to communication can help researchers and teachers isolate pronunciation elements to focus on during instruction. Researchers typically measure listener understanding in two main ways: objective intelligibility measures (e.g., listeners transcribing speech or answering comprehension questions) and/or listeners' rated perceptions of the ease or difficulty of understanding (comprehensibility). The operationalization and measurement of intelligibility is fraught with challenges (Derwing & Munro, 2015), and few researchers have attempted to identify elements of L2 speech linked to intelligibility. However, a fast-emerging area of research is the relationship between the rated comprehensibility of L2 speakers and linguistic features in their speech, with the goal of helping teachers, learners, and language testers isolate and then focus on features that are important for listeners' understanding of L2 speech. For example, Isaacs and Trofimovich (2012) explored correlations between ratings of L2 English speakers' comprehensibility and 19 linguistic measures that focused on characteristics of the L2 speakers' lexis, grammar, discourse, fluency, and pronunciation (individual sounds and prosody). We believe that multiple characteristics, such as L2 speakers' L1 background and their L2 proficiency level, can affect the relative contribution of different aspects of pronunciation to listeners' understanding. In addition, other linguistic dimensions can contribute to listeners' understanding of L2 speech: not only aspects of pronunciation, such as the production of individual sounds or prosody, but also the use of vocabulary and grammar in the L2.

What Pedagogical Approaches Are Effective?

Even with the continuing work to identify appropriate pedagogical norms and targets, the question of effective pedagogical approaches remains. Some older approaches (e.g., the Audio-Lingual Method) were based on behaviorist theoretical frameworks and thus highlighted repetition and imitation of speech models. Many current approaches, such as what Miller (2011) calls the “phonetic approach,” emphasize explicit description of pronunciation patterns and learners’ analysis of speech samples for particular features of pronunciation, as well as the use of communicative activities for pronunciation and fluency practice (Murphy & Baker, 2015). However, these contemporary approaches, which are implemented both in brick-and-mortar classrooms and through computer-enabled software and mobile technology, are rarely grounded in theories of teaching or learning, an issue we return to under the later section, “Pedagogical Approaches.” An exception is the theoretical concept that is regularly cited as a rationale for explicit pronunciation instruction, the concept of noticing as a precursor to learning. As Schmidt (1995) suggested, “what learners notice in input is what becomes intake for learning” (p. 20) and explicit instruction on L2 pronunciation is often justified as a way of helping learners notice formal or functional aspects of L2 pronunciation. However, researchers and practitioners seldom link particular pedagogical approaches to more general theories of learning or specific views of pronunciation development. When pedagogical approaches are not framed in theory, it is difficult to understand which aspects of an approach promote learning or how approaches can be modified for different teaching contexts or learning objectives. If researchers are to rigorously examine the conditions and processes of pronunciation learning, research on pronunciation instruction should be grounded in theoretical views of how pronunciation develops and how L2 learners learn.

How Can Pronunciation Instruction Be Integrated With Other L2 Skills?

In the real world, most learners, particularly young learners, do not have access to stand-alone pronunciation courses. Therefore, for many classroom learners their earliest and most frequent opportunities to be exposed to L2 pronunciation instruction will be in courses that target other areas, such as grammar and vocabulary or reading and writing (Darcy, 2015). Unfortunately, very few studies have explored the effectiveness of pronunciation instruction integrated into a broader L2 course. We agree with Sicola and Darcy (2015), who note that teachers can and should combine teaching pronunciation with the teaching of other L2 skills. However, both formally and informally trained teachers may feel unprepared to integrate pronunciation into regular lessons; teacher education on pronunciation instruction must therefore reframe it not as a specialized activity, but as an activity that can be a fundamental element in teaching an L2.

How Can L2 Speakers Enhance Their Learning Outside L2 Classrooms?

For most learners, the time spent in L2 classrooms, both in the amount and distribution of instructed time, does not promote extensive proficiency development (Muñoz, 2012). This means that learners who wish to enhance their L2 development will benefit from learning activities done outside the classroom, whether the learners are in second language contexts, where L2 exposure is readily available in the physical environment, or in foreign language

contexts, where L2 exposure is more limited and may be primarily available through various media, such as movies or the internet. Learners who can use knowledge and skills developed through instruction to continue learning outside the classroom are more autonomous in their development and less limited by the amount of instructional time. Consequently, learners need to be autonomous in their pronunciation learning outside the classroom, and we maintain that researchers should include measures targeting learning that takes place outside instructional contexts in any study of the effects of L2 pronunciation instruction.

What Is the Role of Teacher Cognition and Education?

All current issues in L2 pronunciation instruction call for teachers with in-depth knowledge of L2 pronunciation, an understanding of the teaching context and learners' goals and expectations, and an ability and confidence to select and carry out learning activities that suit pertinent learning challenges and objectives. Multiple survey and interview studies have revealed that many teachers have low levels of confidence and training in teaching pronunciation (Breitkreutz, Derwing, & Rossiter, 2001; Burns, 2006). In order to improve teacher education for L2 pronunciation, it is important to explore how teachers in preservice and inservice education programs are prepared to teach L2 pronunciation, and how teachers' pedagogical knowledge and skills relate to their pedagogical planning, implementation, and decision-making.

Empirical Evidence

Pronunciation Models and Pedagogical Norms

Learners' and teachers' beliefs about appropriate pronunciation models and pedagogical norms for L2 pronunciation is a widely researched topic, although few studies aim at languages other than English (Drewelow & Theobald, 2007). Findings are tied to particular contexts for language learning and use, as well as specific language ideologies (Litzenberg, 2014), with research often conceptualized within Kachru's (1992) typology of concentric circles of English use. In Inner Circle countries (e.g., Australia, the UK), English has traditionally been the primary language. In Outer Circle countries (e.g., Nigeria), the use of English has typically spread through colonization, so that English is used alongside other languages in educational, commercial, political, and public settings. In Expanding Circle countries (e.g., Poland), English lacks official or historical status.

In Inner Circle countries, learners and teachers often prefer nativelike pronunciation as a pedagogical target or for classroom materials, as opposed to simply intelligible pronunciation. However, learners often cannot articulate their reasons for wanting nativelike pronunciation (Litzenberg, 2014; Scales, Wennerstrom, Richard, & Wu, 2006; Subtirelu, 2013; Young & Walsh, 2010). In Outer Circle contexts, such as Hong Kong, some university students rank native varieties of English more highly than nonnative varieties (Zhang, 2013), while others prefer to use their local accent with a noticeable Cantonese influence for reasons of identity and intelligibility (Sung, 2014). In the little research conducted in Expanding Circle countries in the Middle East, Buckingham (2014) found that although 64% of the 347 Omani university students surveyed wanted to sound like a native English speaker and 72% wanted a native-speaking teacher, they gave favorable ratings to English samples with noticeable Filipino and Arabic influences. In Europe, the majority of the 234 Polish students surveyed expect nativelike English pronunciation to be the model and learning goal

in English classes (Wach, 2011). However, in Scandinavia, half of the 34 surveyed teachers value successful, fluent communication over nativelike English (Ranta, 2010).

It is clear that beliefs about pedagogical norms, particularly for English pronunciation, are associated with multiple factors. The institutional status of English in a given country can be relevant, as learners and teachers in Inner and Expanding Circle countries generally support the use and learning of native English varieties while learners and teachers in Outer Circle countries are often more accepting of other pronunciations. However, other factors, such as participants' teaching experience and the multicultural or multidialectal makeup of the setting (e.g., Hong Kong or Sweden) may also play a role. In choosing pronunciation models and pedagogical norms, many teachers and learners draw on issues relevant to their own contexts, such as the practicality of incorporating particular models in pedagogical materials or their past experience in teaching, learning, or using an L2. These considerations are nontrivial, and the specifics of particular contexts must be acknowledged in any discussion of pronunciation models and norms. If a restricted set of models or norms is advocated for all teaching and learning contexts, it sends a message to teachers and learners that their own concerns and conditions are irrelevant.

Pedagogical Targets Linked to Interlocutor Understanding

A growing area of research investigates which linguistic elements of L2 speech are linked to interlocutor difficulties in understanding, with the idea that those elements can serve as pedagogical targets if learners' primary goal is to make their speech understandable. Intelligibility is a challenging construct to measure (Derwing & Munro, 2015); therefore, most research on intelligibility has relied on interlocutor behavior (signs of communication breakdowns or transcriptions of speech) to identify speaker difficulties. For instance, L2 English speech by L1 Japanese learners can be problematic to L1 English listeners due to deletion of consonants, misplaced word stress, devoicing of consonants, vowel substitutions, and substitution of /ɾ/ (flap) for /l/ (Suenobu, Kanzaki, & Yamane, 1992). For L1 English and L1 Cantonese listeners, difficulties in understanding L2 English speakers from a variety of Asian backgrounds are related to misplaced word stress and substitutions of initial consonants and vowels (Sewell, 2015; Zielinski, 2008). And similar features, along with deletion of segments in consonant clusters and misplaced nuclear (phrase) stress, can impair intelligibility between L1 speakers of Asian languages (Deterding & Kirkpatrick, 2006; Matsumoto, 2011). Through analyses of the mutual intelligibility of L2 English speakers from different L1s, Jenkins (2000) identified several linguistic elements as crucial for intelligibility, including production of consonants and consonant clusters, aspiration of stops, vowel length, and nuclear stress. Later research supported these results, but nontarget placement of word stress and vowel substitutions were also implicated in difficulties in understanding (Kennedy, 2012; O'Neal, 2015).

With respect to L2 comprehensibility (rated ease of understanding), Isaacs, Saito, Trofimovich, and their colleagues have shown that L2 English comprehensibility ratings for different L1 speaker groups (French, Farsi, Mandarin, Japanese) are broadly associated with two dimensions: pronunciation (individual segments, prosody, fluency) and lexicogrammar (varied/appropriate use of words and accurate/complex grammar) (Saito, Trofimovich, & Isaacs, 2015, 2016; Saito, Webb, Trofimovich, & Isaacs, 2015). However, linguistic links to comprehensibility also depend on speakers' L1 background and the speaking task (Crowther, Trofimovich, Isaacs, & Saito, 2015; Crowther, Trofimovich, Saito, & Isaacs, 2015). For example, while segmental errors were linked to L1 Chinese speakers' comprehensibility, it

was lexicogrammar that contributed to comprehensibility for L1 Hindi-Urdu speakers. The only non-English study in this work has shown that comprehensibility of L1 English learners of L2 German was predicted by fluency measures and by measures of accuracy for lexis, morphology, stress placement, and the pronunciation of segments specific to the German but not the English phonological inventory (O'Brien, 2014).

Clearly, listeners' perceived and actual understanding of L2 speech is linked not only to pronunciation-related elements but to many other language- and discourse-related features (see also Kennedy et al., 2015). Several speech elements linked to understanding, such as consonant and vowel substitutions, syllable structure, and nuclear stress, cut across different contexts and speaker groups; however, other difficulties in understanding are related to specific features for particular L1 groups. Therefore, identifying pedagogical approaches that are relevant and effective for learners from multiple language backgrounds is no easy task.

Efficacy of Pronunciation Instruction

Several recent research reviews and a meta-analysis have shown that learners' pronunciation improves for both individual segment and prosody targets after receiving pronunciation instruction (Lee, Jang, & Plonsky, 2015; Saito, 2012; Thomson & Derwing, 2015). This is true both for instruction which includes a communicative focus, and for instruction which targets only formal aspects of pronunciation. Lee et al.'s meta-analysis has shown that instructional effects are stronger for longer interventions and interventions that include corrective feedback. Instruction in L2 settings shows stronger effects than instruction in contexts where the target language is not readily available outside the classroom, and effects are stronger for learners receiving laboratory-based than for those receiving classroom-based instruction. Interventions delivered by humans have stronger effects than interventions that solely or in part use technology. In addition, stronger effects are seen for restricted outcome measures, such as reading aloud, than for more open-ended measures, such as picture descriptions.

Computer-Aided Pronunciation Teaching (CAPT)

In much computer-aided pronunciation teaching (CAPT) research, pronunciation is modeled by manipulating the frequency and/or salience of targeted pronunciation features, with learners often receiving feedback on their productions; we illustrate three strands of this work, all using freely available technology. In the first strand, visual information about pronunciation is often included in instructional materials to raise learners' awareness about targeted features or to provide unambiguous feedback about their pronunciation. For example, to focus on the pronunciation of geminate consonants in Japanese, Motohashi-Saigo and Hardison (2009) presented target words to L2 Japanese learners under two conditions: oral only and oral paired with visual speech waveform displays in Praat. Praat is a freely downloadable program for speech analysis and editing, which requires some linguistic and technical expertise (Boersma & Weenink, 2013). After each presented word, learners selected one of three written response options, receiving immediate feedback on the accuracy of their selection. Learners in both conditions significantly improved in their production of geminate consonants but learners in the visual condition improved to a greater degree.

The second strand incorporates the use of high-variability phonetic training (HVPT), which has a long history in speech research; it is based on theoretical views of how humans form categories and learn from examples. In HVPT research, learners are typically presented

with multiple tokens of L2 pronunciation targets, such as vowels or consonants, spoken by different voices, with learners receiving feedback on their accuracy after perceiving or producing targets. Until recently, most HPVPT research has been conducted in laboratories because of the challenge of creating diverse sets of materials and presenting them to learners. Thomson (2011) used speech-presentation software, along with Praat, to develop and test an HVPT tool to teach English vowels, reporting a significant improvement in the intelligibility of vowel production by L1 Mandarin learners after instruction. Thomson later converted this tool to a web-based interface, which offers free HVPT training for North American English vowels and consonants (<http://www.englishaccentcoach.com>).

In the third strand, researchers have used internet-enabled technologies for learners to practice and give each other feedback on pronunciation. Ducate and Lomicka (2009) found little improvement in the comprehensibility of university students in intermediate-level French and German courses after they completed eight podcasts and commented on the content of classmates' podcasts, while university students in a Spanish phonetics course did improve in accent ratings after completing and providing constructive, pronunciation-related feedback for six podcasts (Lord, 2008). Drawing on the Interaction Hypothesis, Bueno Alastuey (2010) paired L1 Spanish students with interaction partners whose L1s were Spanish, Turkish, or English. The pairs engaged in six separate interactions in English via a synchronous voice-mediated communication tool (e.g., Skype). Ratings of the students' pronunciation significantly improved over time, no matter the partner's L1. Bueno Alastuey suggests that the improvement may have resulted from the increased speaking time compared to the time available in students' regular classrooms, as well as the potential for individualized practice and negotiation of meaning leading to modified pronunciation output. Although Lee et al.'s (2015) meta-analysis of past research showed that interventions by humans have stronger instructional effects than interventions that use technology, it appears that CAPT holds great potential in addressing learners' pronunciation needs. However, much existing research draws on technology that requires equipment, financial resources, or technical expertise that may not be readily available to teachers; this research thus has little relevance for many instructional contexts, especially those with little institutional support for technology.

Pedagogical Approaches

Until the end of the 20th century, research on L2 pronunciation instruction mainly focused on whether learners showed any pronunciation differences after receiving instruction, or whether an emphasis on certain targets, compared to others, resulted in improved pronunciation. Most instruction was described atheoretically, with few clear links to social, cultural, or psychological principles of language learning and teaching. There is, however, current research that is clearly set in theoretical frameworks. Couper (2011) used an approach based on cognitive phonology to investigate how language can be used as a tool to socially construct meaning and achieve communication. Cognitive phonology frames pronunciation as a tool in a meaning-making process, with speakers' concepts of sounds playing an important role in how they use the tool of pronunciation to express meaning (Fraser, 1999). In Couper's study, adult learners of English received four brief lessons targeting syllable endings, with each lesson lasting under an hour. All groups engaged in listening and speaking practice, but two groups focused on identifying and receiving feedback on words whose syllable endings could be confused, and two groups worked to create (socially construct) their own metalanguage (learner-friendly description) about English syllable endings. For example, corrective feedback on learners' pronunciation of "a drunk snail" might be "make the 'k'

quieter and shorter.” After instruction, the groups who listened to confusable syllable endings were significantly better at perceiving syllable endings than the other groups, and the groups who created their own metalanguage were significantly better at producing syllable endings than the other groups.

In other theoretically driven work, Saito and his colleagues have targeted the construct of noticing, highlighting the important role of form-focused instruction (FFI) in pronunciation teaching. FFI refers to techniques that “draw attention to target language features that learners would otherwise not use or even notice in communicatively oriented classroom input” (Saito & Lyster, 2012, p. 596). In Saito and Lyster’s study, instructional materials targeting English /r/ featured thematically focused and meaningful tasks with structured, typographically enhanced input. After 4 hours of FFI instruction in Canada, L1 Japanese learners who also received corrective feedback in the form of recasts (teacher’s reformulations of learners’ nontarget utterances) showed significant improvement in the intelligibility of their /r/ production while those who received only FFI did not show improvement. In a similar study in Japan, Saito (2013) showed that L1 Japanese learners receiving FFI with corrective feedback moderately improved in /r/ production only for familiar lexical items, but learners who also received explicit instruction before completing focused tasks showed large improvement for both familiar and unfamiliar lexical items. Saito (2015) also reported a link between corrective feedback and L2 pronunciation, such that higher numbers of recasts provided to learners were associated with greater accuracy gains in their English /r/ production.

Integration of Pronunciation in L2 Instruction

In the world of research, pronunciation instruction typically takes place in courses devoted to L2 oral skills. However, Darcy, Ewert, and Lidster (2012) note that many teachers cite lack of time as a reason for not teaching pronunciation. One solution is to integrate pronunciation instruction into regular classes. Although several researchers have described ways of integrating pronunciation into daily lessons or with other language skills (Chela-Flores, 2001; Nicolaidis & Mattheoudakis, 2012), little research has targeted the effectiveness of such integration. One exception is Roccamo (2015), who incorporated 10-minute pronunciation modules on four pronunciation features into a four-skill (reading, writing, speaking, listening) beginner-level university German course, offered four times weekly. The modules featured perception and production activities and peer feedback. Both the treatment and comparison groups significantly improved in overall comprehensibility over 8 weeks, but the treatment group significantly improved for three of the four features in two read-aloud tasks while the control improved only for two features in one read-aloud task. These results show the possibility and benefits of integration, even at beginner levels.

Learner Autonomy

Autonomy refers to learners’ self-initiated behaviors and actions whose purpose is to further the learning and use of L2 pronunciation, and researchers are now increasingly turning to the issue of how L2 pronunciation development is linked to strategy instruction and out-of-class exposure, which can promote learner autonomy by allowing learners to use an L2 for authentic purposes (Moyer, 2011). In one study (Sardegna & McGregor, 2013), L2 English graduate students who received strategy instruction in a university pronunciation course improved in their pronunciation of targeted aspects of prosody, and when recorded 5–25 months after the course, those students who self-reported appropriate and consistent use

of the strategies after the course showed a less dramatic drop in their pronunciation scores than students who reported minimal or inappropriate strategy use. The strategy instruction included teacher-guided instruction, modelling, and practice of specific elements of pronunciation in class, students' out-of-class practice and monitoring of those elements, and students' reflections on the teacher's feedback on students' out-of-class recordings. In terms of out-of-class L2 exposure, multiple studies have shown that learners' quantity or quality of L2 exposure outside the classroom is linked to the quality of their pronunciation (e.g., Kennedy & Trofimovich, 2010). However, apart from study-abroad experiences, purposeful interventions that require learners to engage in L2 use outside the classroom have not been well documented in research on pronunciation instruction. A focus on learner awareness of pronunciation (in terms of their conscious recognition and understanding of various aspects of speech, as they are used in a target language) and learner autonomy beyond the classroom is also warranted because L2 speakers are often unaware of their pronunciation difficulties, with many overestimating the degree to which they are understood by their interlocutors (Trofimovich, Isaacs, Kennedy, Saito, & Crowther, 2016).

Teacher Education

Early survey studies of L2 teachers, predominantly teachers of English in second language contexts, showed a worrying trend. Many, if not most, of the surveyed teachers lacked confidence in their knowledge of L2 pronunciation, focused on pronunciation only rarely in their teaching, and resorted to a limited range of activities when they targeted pronunciation (Breitkreutz et al., 2001; Burns, 2006). In recent research, teachers reported more confidence and training opportunities but similar types of activities were used (Foote, Holtby, & Derwing, 2011). In foreign language contexts, surveys conducted in Europe and Brazil show a high incidence of formal teacher education, and a range of experience with learning the target language sound system, but little formal training in teaching pronunciation (Buss, 2015; Henderson et al., 2012). Research in both second and foreign language contexts has also revealed that some teachers who identify as nonnative speakers of the L2 lack confidence in their pronunciation teaching ability because of their own inability to provide a model of native pronunciation (Reis, 2011; Tum, 2013).

However, as Baker and Murphy (2011) note, very little is known about what teachers actually do in the classroom, especially as related to teacher cognitions about pronunciation instruction. Tergujeff (2012), who observed four teachers of English in Finnish primary and secondary schools over 1 week, found that teachers generally used traditional teacher-centered teaching activities, such as explicit metalinguistic instruction, listening and repeating, and reading aloud, and that the main pedagogical targets were segments typically problematic for L1 Finnish learners. In another study targeting pronunciation teaching practices, Foote, Trofimovich, Collins, and Soler Urzúa (2016) analyzed a corpus of 40 hours of videotaped lessons for 11- to 12-year-old learners in an intensive English program in Quebec, Canada. Pronunciation accounted for only 10% of all language-related episodes, generally taking the form of teachers' corrective feedback on individual segments.

While pronunciation teaching practices can vary greatly across contexts, influenced by the local pedagogical culture and objectives, teachers' practices can also be determined by their current and past beliefs (Baker & Murphy, 2011). Baker (2014) investigated the teaching practices and cognitions of five teachers of oral communication in a North American intensive ESL program. Teacher beliefs, attributed to teachers' previous education in pronunciation pedagogy, clustered around three main areas: learners' perception of L2 speech

Sara Kennedy and Pavel Trofimovich

is important for their pronunciation, kinesthetic/tactile activities, which link particular pronunciation features to particular physical movements like clapping, are crucial for learners to improve, and pronunciation instruction can be boring, which was reflected in the search by some teachers for extra material. Burri (2015) explored the influence of teacher education by tracking the cognitions of teachers taking a graduate-level pronunciation pedagogy course at an Australian university. Changes in cognitions were individual to each teacher, but all teachers showed increased awareness of the nature and importance of prosody in English. For those who were L2 speakers, this change was linked to their self-perceived improvement in English pronunciation; teachers who were L1 English speakers also became more aware that L2 English speakers could be effective teachers of pronunciation. In sum, there is clearly room for further research into effective pedagogical and professional development practices for preservice and inservice teachers, with the aim of both encouraging and sustaining classroom teaching of L2 pronunciation that is reflective, informed, and suits the instructional conditions and objectives.

Pedagogical Implications

Pronunciation Models, Pedagogical Norms, and Targets

As Rogerson-Revell (2011) points out, pronunciation models are not the same as pedagogical norms (learning and teaching goals). In many contexts, most or all available language teaching and learning materials may be based on a native variety (a pronunciation model). Teachers and learners might rely on this model as an overall outline of L2 phonology, either because no other pronunciation models are available, because of the model's prestige value, or because it provides consistent reference points for teaching, learning, and assessment. However, there is *no inherent need* for the pedagogical norms and targets selected by administrators, teachers, and learners to closely correspond to the pronunciation model. Decisions about appropriate pedagogical norms and targets should be shaped by the context of language teaching, learning, and use, both within the classroom and outside it. For example, teachers targeting spoken language varieties as objects of academic study (e.g., in a university phonetics course) will probably aim for learners to use pronunciation that corresponds to that variety. In other contexts, however, the wider sociopolitical climate (e.g., bilingualism, as in Hong Kong, or unequal power status for languages used in postcolonial contexts, such as Rwanda) may mean that some teachers and learners do not want to adopt a native variety as a pedagogical norm. Therefore, pronunciation models and pedagogical norms and targets need to be appropriate to the specific contexts and conditions of language teaching, learning, and use.

Listener Understanding as a Pedagogical Norm and Target

Many children and most adults who are taught L2 pronunciation will not attain nativelike pronunciation, so it seems unreasonable and impractical to adopt native norms for most teaching contexts. Going beyond native norms, pronunciation specialists have put forward different sets of factors for teachers to consider in selecting pedagogical targets. Many of these relate to the intelligibility or frequency of particular sounds or words in the spoken L2 or to the role of particular sounds or sound patterns in listeners' processing of speech (Darcy et al., 2012; Munro & Derwing, 2006; Rogerson-Revell, 2011). Listener understanding is a crucial component in these factors. Research-based findings about pronunciation features

linked to listeners' understanding are still emerging, especially for L2s other than English and for different L1 speakers of a given L2. However, an important complement to research findings is experiential knowledge, both from teachers and learners, who discover features important for understanding simply from the experience of observing or doing oral communication in the L2. This knowledge is sometimes formally organized and disseminated (e.g., Swan & Smith, 2001), but most typically spreads among teachers and learners during classes and informal discussions. In any case, whatever the teaching context, listener understanding is an important component of pedagogical norms and targets.

Influences on Listeners' Understanding

Although listeners' understanding may be important for pedagogical norms in pronunciation instruction, pronunciation is only one of the elements contributing to communication success. Lexical, grammatical, and discourse characteristics of L2 speech, task characteristics, and listeners' attitude and previous L2 exposure can all influence listener understanding. This means that learners and teachers who want to work on intelligibility must be open to working on more than simply pronunciation, because listeners' understanding is influenced by much more than pronunciation.

Going Beyond Explicit Instruction

Although pronunciation instruction can be effective, there is insufficient research evidence to support the effectiveness of particular pedagogical approaches. Leaving out specific approaches, there is evidence that learners' pronunciation can improve without explicit instruction if tasks are communicative but include a focus on form and consistent, targeted corrective feedback. Similarly, learners' participation in socially constructing metalanguage about pedagogical targets can lead to significant improvement in pronunciation. Teachers should carefully consider the degree to which their learners need explicit, teacher-directed instruction as opposed to other types of awareness-raising or practice opportunities.

Engaging Learners in Learning

Learners do not react to instruction in identical ways; their motivation, attitude, L2 exposure, and many other individual characteristics can promote or hinder their learning. Teachers can address some of these characteristics, for instance, through class discussion of learners' attitudes to pronunciation or through tasks promoting learner autonomy. When teachers and learners understand how these characteristics are related to teaching and learning, it will be easier to understand how different pedagogical activities can be more or less effective for learners, and learners will be better prepared to contribute to their own learning.

Teaching Teachers to Integrate Pronunciation

The most frequent opportunities for classroom teaching of pronunciation lie not in supplementary teaching, but in teaching that is a regular part of the class. This situation means that teacher education in pronunciation teaching methodology needs to focus not on developing or using standalone pronunciation activities, but on exploiting opportunities for pronunciation teaching using existing classroom materials and activities, such as vocabulary learning,

Sara Kennedy and Pavel Trofimovich

grammar practice, or role plays. Although challenging to accomplish, this integration will allow teachers to draw on familiar materials and activities in order to increase or enhance their pronunciation teaching. Learners will be using pronunciation in habitual contexts, supporting the transfer of learning to learners' spontaneous speech. Pronunciation-related teacher education should therefore focus on integration, not specialization.

Teaching Tips

- Help learners to perceive L2 sounds as well as to produce them. Learners who struggle to accurately perceive L2 segments or prosody also struggle to accurately produce them.
- Use tasks that are meaningful but also promote repetition of target forms. Pedagogical findings show that, in addition to explicit instruction, meaningful, form-focused tasks that use familiar vocabulary can help learners improve their pronunciation.
- Teach learners to use all their linguistic resources in order to get their message across. Listeners' understanding of L2 speech is affected by pronunciation but also by vocabulary, grammar, and discourse organization.
- Exploit existing materials to teach and learn pronunciation so that instruction can be integrated into existing classes and learners can encounter familiar and meaningful texts.

Future Directions

The reviews and meta-analyses of pronunciation instruction (Lee et al., 2015; Thomson & Derwing, 2015) noted several overall weaknesses in the sampling, design, and reporting of pronunciation instruction. Researchers were urged to recruit larger participant samples, with a wider variety of target languages and ages, and to include more open-ended outcome measures, which better represent authentic use of spontaneous speech, and more delayed posttests in order to explore the resilience of instructional effects. Finally, researchers were advised to more clearly describe the nature of the pronunciation instruction provided. These are constructive and valuable suggestions that, if followed, will expand the scope of pronunciation research and will help make a stronger case for the significance of particular findings for the real world.

With respect to pedagogy, we offer other recommendations. Because the majority of opportunities for pronunciation teaching are in language classes that are not devoted to pronunciation, more observational research is needed to understand what sort of pronunciation teaching occurs in these classes. Some researchers might be concerned about spending time and resources on observation, only to find that very little pronunciation teaching is happening. One possibility to mitigate such concerns is to recruit teachers to conduct action research in their own classes, logging their actual or potential pronunciation teaching. Another possibility is to incorporate formal observation of classes in teacher education programs, so that preservice teachers are creating corpora of classroom teaching as part of their coursework.

In addition to data on classroom teaching, research on teachers' cognition and their classroom practices is crucial. Teachers have reasons for the decisions they make before, during, and after they teach. Understanding how teachers' practices are linked to their cognitions can highlight critical areas for teacher education. Learners should also be asked for their views

on teachers' observed classroom practices, as it is learners who are most directly affected by these practices. The most obvious environments for this research are classes that do not target pronunciation. Finally, although there are several reports on teacher education initiatives for pronunciation pedagogy (e.g., Burns, 2006), to date only Burri (2015) has explored teachers' developing cognitions in teacher education, and no research on the effects of teacher education on actual pronunciation teaching practices has been published. To develop effective teacher education, it is essential to observe its results on what teachers do.

References

- Abercrombie, D. (1949). Teaching pronunciation. *ELT Journal*, 3, 113–122.
- Abrahamsson, N., & Hylltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249–306.
- Baker, A. (2014). Exploring teachers' knowledge of second language pronunciation techniques: Teacher cognitions, observed classroom practices, and student perceptions. *TESOL Quarterly*, 48, 136–163.
- Baker, A., & Murphy, J. (2011). Knowledge base of pronunciation teaching: Staking out the territory. *TESL Canada Journal*, 28, 29–50.
- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 13–34). Amsterdam: John Benjamins.
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.52) [Computer Software]. Retrieved from <http://www.praat.org>
- Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19, 51–61.
- Buckingham, L. (2014). Attitudes to English teachers' accents in the Gulf. *International Journal of Applied Linguistics*, 24, 50–73.
- Bueno Alastuey, M. C. (2010). Synchronous-voice computer-mediated communication: Effects on pronunciation. *CALICO Journal*, 28, 1–20.
- Burns, A. (2006). Integrating research and professional development on pronunciation teaching in a national adult ESL program. *TESL Reporter*, 39, 34–41.
- Burri, M. (2015). Student teachers' cognition about L2 pronunciation instruction: A case study. *Australian Journal of Teacher Education*, 40(10), 66–87.
- Buss, L. (2015). Beliefs and practices of Brazilian EFL teachers regarding pronunciation. *Language Teaching Research*, 20, 1–19.
- Chela-Flores, B. (2001). Pronunciation and language learning: An integrative approach. *International Review of Applied Linguistics in Language Teaching*, 39, 85–101.
- Cook, V. J. (1992). Evidence for multicompetence. *Language Learning*, 42, 557–591.
- Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: Socially constructed metalanguage and critical listening. *Language Awareness*, 20, 159–182.
- Crowther, D., Trofimovich, P., & Isaacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility: Nonnative listeners' perspectives. *Journal of Second Language Pronunciation*, 2, 160–182.
- Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2015). Does a speaking task affect second language comprehensibility? *Modern Language Journal*, 99, 80–95.
- Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2015). Second language comprehensibility revisited: Investigating the effects of learner background. *TESOL Quarterly*, 49, 814–837.
- Darcy, I. (2015). Powerful and effective pronunciation instruction: How can we achieve it? [PDF document]. Retrieved from http://www.upf.edu/masterlinguistica/_pdf/Pronunciation_Instruction_Darcy_June_2015_UPF.pdf

Sara Kennedy and Pavel Trofimovich

- Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox." In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd annual pronunciation in second language learning and teaching conference* (pp. 93–108). Ames, IA: Iowa State University.
- de Bot, K., & Mailfert, K. (1982). The teaching of intonation: Fundamental research and classroom applications. *TESOL Quarterly*, 16, 71–77.
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 20, 1–16.
- Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, 63, 163–185.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Deterding, D., & Kirkpatrick, A. (2006). Emerging South-East Asian Englishes and intelligibility. *World Englishes*, 25, 391–409.
- Drewelow, I., & Theobald, A. (2007). A comparison of the attitudes of learners, instructors, and native French speakers about the pronunciation of French: An exploratory study. *Foreign Language Annals*, 40, 491–520.
- Ducate, L., & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning & Technology*, 13, 66–86.
- Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Berlin: Mouton de Gruyter.
- Foote, J. A., Holtby, A., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, 29, 1–22.
- Foote, J. A., Trofimovich, P., Collins, L., & Soler Urzúa, F. (2016). Pronunciation teaching practices in communicative ESL classes. *Language Learning Journal*, 44, 181–196.
- Fraser, H. (1999). ESL pronunciation teaching: Could it be more effective? *Australian Language Matters*, 7, 7–8.
- Fraser, H. (2004). Constraining abstractness: Phonological representation in the light of color terms. *Cognitive Linguistics*, 15, 239–288.
- Gilbert, J. (2001). Six pronunciation priorities for the beginning student. *CATESOL Journal*, 13, 173–182.
- González-Bueno, M. (1997). The effects of formal instruction on the acquisition of Spanish stop consonants. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish* (Vol. 2, pp. 57–75). Somerville, MA: Cascadia Press.
- Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U., & Curnick, L. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, 10, 5–27.
- Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility. *Studies in Second Language Acquisition*, 34(3), 475–505.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
- Kachru, B. B. (1992). *The other tongue: English across cultures*. Urbana, IL: University of Illinois Press.
- Kang, O. (2015). Learners' perceptions toward pronunciation instruction in three circles of World Englishes. *TESOL Journal*, 6(1), 59–80.
- Kennedy, S. (2012). When non-native speakers misunderstand each other: Identifying important aspects of pronunciation. *Contact*, 38, 49–62.
- Kennedy, S., Guénette, D., Murphy, J., & Allard, S. (2015). Le rôle de la prononciation dans l'intercompréhension entre locuteurs de français lingua franca. *Canadian Modern Language Review*, 71, 1–25.
- Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness*, 19, 171–185.

- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal*, 73, 440–464.
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36, 345–366.
- Levis, J.M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39, 369–377.
- Litzenberg, J. (2014). Pre-service teacher perspectives towards pedagogical uses of non-native and native speech samples. *International Journal of Applied Linguistics*, 26, 1–22.
- Lord, G. (2008). Podcasting communities and second language pronunciation. *Foreign Language Annals*, 41, 364–379.
- Matsumoto, Y. (2011). Successful ELF communications and implications for ELT: Sequential analysis of ELF pronunciation negotiation strategies. *Modern Language Journal*, 95, 97–114.
- Miller, J. (2011). Teaching pronunciation with phonetics in a beginner French course: Impact on sound perception. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th pronunciation in second language learning and teaching conference* (pp. 109–123). Ames, IA: Iowa State University.
- Motohashi-Saigo, M., & Hardison, D.M. (2009). Acquisition of L2 Japanese geminates: Training with waveform displays. *Language Learning & Technology*, 13, 29–47.
- Moyer, A. (2011). An investigation of experience in L2 phonology: Does quality matter more than quantity? *Canadian Modern Language Review*, 67, 191–216.
- Muñoz, C. (2012). The significance of intensive exposure as a turning point in learners' histories. In C. Muñoz (Ed.), *Intensive exposure experiences in second language learning* (pp. 141–160). Bristol: Multilingual Matters.
- Munro, M.J., & Derwing, T.M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73–97.
- Munro, M.J., & Derwing, T.M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34, 520–531.
- Munro, M.J., & Derwing, T.M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58, 479–502.
- Munro, M.J., & Derwing, T.M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, 44, 316–327.
- Murphy, J., & Baker, A. (2015). History of pronunciation teaching. In M. Reed & J.M. Levis (Eds.), *The handbook of English pronunciation* (pp. 35–65). Hoboken, NJ: Wiley.
- Nicolaidis, K., & Mattheoudakis, M. (2012). The PRO-VOC method: Combining pronunciation and vocabulary. *International Review of Applied Linguistics in Language Teaching*, 50, 303–321.
- O'Brien, M.G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning*, 64, 715–748.
- O'Neal, G. (2015). Segmental repair and interactional intelligibility: The relationship between consonant deletion, consonant insertion, and pronunciation intelligibility in English as a lingua franca in Japan. *Journal of Pragmatics*, 85, 122–134.
- Purcell, E.T., & Suter, R.W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, 30, 271–287.
- Ranta, E. (2010). English in the real world vs. English at school: Finnish English teachers' and students' views. *International Journal of Applied Linguistics*, 20, 156–177.
- Reis, D.S. (2011). Non-native English-speaking teachers (NNESTs) and professional legitimacy: A sociocultural theoretical perspective on identity transformation. *International Journal of the Sociology of Language*, 208, 139–160.
- Roccamo, A. (2015). Teaching pronunciation in just ten minutes a day: A method for pronunciation instruction in first-semester German language classrooms. *Die Unterrichtspraxis*, 48, 59–83.
- Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. London: Bloomsbury.
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 46, 842–854.

Sara Kennedy and Pavel Trofimovich

- Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /ɪ/ to adult Japanese learners. *Language Learning*, 63, 499–529.
- Saito, K. (2015). Variables affecting the effects of recasts on L2 pronunciation development. *Language Teaching Research*, 19, 276–300.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɪ/ by Japanese learners of English. *Language Learning*, 62, 595–633.
- Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgements to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*. Published online 29 September 2015. doi:10.1093/applin/amv047
- Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37, 217–240.
- Saito, K., Webb, S., Trofimovich, P., & Isaacs, T. (2015). Lexical profiles of comprehensible second language speech: The role of appropriateness, fluency, variation, sophistication, abstractness and sense relations. *Studies in Second Language Acquisition*, 38, 677–701.
- Sardegna, V.G., & McGregor, A. (2013). Scaffolding students' self-regulated efforts for effective pronunciation practice. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th pronunciation in second language learning and teaching conference* (pp. 182–193). Ames, IA: Iowa State University.
- Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of accent. *TESOL Quarterly*, 40, 715–738.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: University of Hawai'i Press.
- Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford: Oxford University Press.
- Sewell, A. (2015). The intranational intelligibility of Hong Kong English accents. *System*, 49, 86–97.
- Sicola, L., & Darcy, I. (2015). Integrating pronunciation into the language classroom. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation* (pp. 467–483). Malden, MA: Wiley.
- Sisson, C.R. (1970). The effect of delayed comparison in the language laboratory on phoneme discrimination and pronunciation accuracy. *Language Learning*, 20, 69–88.
- Subtirelu, N. (2013). What (do) learners want (?): A re-examination of the issue of learner preferences regarding the use of “native” speaker norms in English language teaching. *Language Awareness*, 22, 270–291.
- Suenobu, M., Kanzaki, K., & Yamane, S. (1992). An experimental study of intelligibility of Japanese English. *International Review of Applied Linguistics in Language Teaching*, 30, 146–156.
- Sung, C.C.M. (2014). Accent and identity: Exploring the perceptions among bilingual speakers of English as a lingua franca in Hong Kong. *International Journal of Bilingual Education and Bilingualism*, 17, 544–557.
- Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge: Cambridge University Press.
- Sweet, H. (1900). *The history of language*. London: Dent.
- Tergujeff, E. (2012). English pronunciation teaching: Four case studies from Finland. *Journal of Language Teaching and Research*, 3, 599–607.
- Thomson, R.I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Calico Journal*, 28, 744–765.
- Thomson, R.I., & Derwing, T.M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36, 326–344.
- Trofimovich, P., Isaacs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: Investigating self- and other-perception of second language speech. *Bilingualism: Language and Cognition*, 19(1), 122–140.

- Tum, D. O. (2013). *A study of non-native student teachers' feelings of foreign language teacher anxiety* (Doctoral dissertation). Institute of Education, University of London.
- Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, 4, 114–136.
- Wach, A. (2011). Native-speaker and English as a lingua franca pronunciation norms: English majors' views. *Studies in Second Language Learning and Teaching*, 1, 247–266.
- Young, T. J., & Walsh, S. (2010). Which English? Whose English? An investigation of “non-native” teachers' beliefs about target varieties. *Language, Culture, and Curriculum*, 23, 123–137.
- Zhang, Q. (2013). The attitudes of Hong Kong students towards Hong Kong English and Mandarin-accented English. *English Today*, 29, 9–16.
- Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69–84.

Vocabulary Acquisition

Beatriz González-Fernández and Norbert Schmitt

Background

Vocabulary is an essential component of any language, and thus it is a critical part of second language (L2) acquisition (e.g., Nation, 2013; Willis & Ohashi, 2012). Vocabulary knowledge influences both productive skills (speaking and writing) and receptive skills (reading and listening), and is considered a key predictor of general language proficiency (Alderson, 2007; Laufer & Goldstein, 2004). L2 learners often acknowledge that the lack of or poor vocabulary knowledge is the main reason for their difficulties in acquiring, comprehending, and using a L2 (Nation, 2013). This chapter will focus on the key principles of vocabulary acquisition and how they guide current vocabulary pedagogy. Some of these issues include the overall inattention to vocabulary instruction during different eras, the importance of learning a large number of words, the necessity of learning various aspects about these words, receptive and productive mastery, knowledge of formulaic language, the incremental nature of vocabulary acquisition, and the need for multiple incidental and intentional exposures to a word in order to develop a proficient enough mastery to be able to use it appropriately in all situations.

Key Concept

Vocabulary acquisition: All the processes involved in learning lexical items (i.e., single words and formulaic language) in sufficient depth to be able to use them both productively and receptively, by means of multiple incidental and intentional encounters with these items in varied contexts.

Historical Background

Grammar Translation Method

The grammar translation method dominated from the end of the 18th century, all the way throughout the 19th and 20th centuries, and is still used in many foreign language teaching contexts today. The focus of instruction was mainly grammar, and vocabulary was largely

disregarded, attended to in the form of bilingual lists of archaic words to be used in the translation of literary texts (Zimmerman, 1997).

Vocabulary Control Movement (VCM)

The early 20th century was characterized by the vocabulary control movement (VCM) (especially in the British sphere of influence), which attempted to raise the status of vocabulary in L2 learning. For the first time, vocabulary was considered the crucial element in language teaching. Similar to the grammar translation method, the VCM was based on the use of vocabulary lists. However, unlike in the previous period, during the VCM researchers focused on using innovative and systematic criteria to select the most useful vocabulary for language learning, such as the use of word frequency. The most famous list derived from this movement was the *General Service List* (GSL) (West, 1953), which presented the most useful 2,000 words of English.

Audio-lingual Method and Chomsky

In America, the audio-lingual method was developed during World War II, with a rationale based on behaviorism. The main focus of this method was the acquisition of grammatical patterns through repetition, and the acquisition of vocabulary was downplayed. Therefore, only a very few simple and familiar words were explicitly taught, as it was assumed that vocabulary would be picked up incidentally through exposure to the language without the need for explicit instruction (Zimmerman, 1997). Subsequently, Chomsky's (1957) views shifted the field's theoretical understanding of language acquisition, but his notion of Universal Grammar did not change the relative neglect of vocabulary pedagogically, and the VCM that was taking place in Britain at the time was largely ignored.

Communicative Language Teaching

In the communicative language teaching method (1970s), language teaching focused on the acquisition of functional language (e.g., how to make a request, how to apologize), and the focus changed from using grammar accurately to using the L2 fluently and appropriately in real, meaningful communication, where the attention was on the message (Larsen-Freeman, 2000). Despite this meaning-based, communicative approach, however, once again vocabulary occupied a secondary place in language teaching. Vocabulary items were thought to be acquired incidentally by exposure, without the need of explicit instruction, and thus there was a lack of a principled approach for vocabulary teaching.

Reemergence of Vocabulary

In 1980, Paul Meara highlighted the striking neglect of vocabulary acquisition as part of L2 learning, despite its crucial importance for language use. Indeed, around the time of Meara's observation, there began an increasing emphasis on the role of vocabulary in language teaching, and some researchers started to draw attention to the need of studying the processes of vocabulary acquisition (e.g., Levenston, 1979; Richards, 1976). However, it was not until 1990 when Paul Nation provided the key impetus to study vocabulary, with his book *Teaching and Learning Vocabulary*, which nearly singlehandedly inspired a renewed interest both in vocabulary research and teaching. He proposed for the first time a principled, systematic

approach to vocabulary instruction, bringing back some of the ideas of the VCM. He argued that a frequency-based approach is the best way of selecting and organizing the vocabulary to be taught in a language course, hence emphasizing the value of corpus studies.

The New Millennium

Vocabulary acquisition now has a central role in the field of Instructed Second Language Acquisition (ISLA), and there has been an explosion in the amount of vocabulary research taking place. Nation, writing in 2013, estimated that over 30% of all the research on vocabulary since 1900 was published in the previous 11 years. This wealth of research has been very informative, and the following section will distill some of the major insights gained about vocabulary knowledge and acquisition.

To sum up, this overview shows how despite the importance of vocabulary, its role in language instruction has an uneven past, being undervalued/disregarded at some points in time and emphasized at others.

Current Issues and Empirical Evidence

The Nature of the Lexicon

Research has found that L2 learners' vocabulary knowledge comprises not only knowing a multitude of words, but also gaining various types of knowledge about each word, and establishing connections between multiple lexical items to create semantic networks (Cremer, Dingshoff, Beer, & Schoonen, 2010). However, it is still unclear how vocabulary is stored and processed in the mental lexicon. It is known that words are not unrelated and independent from each other, but rather they are linked in multiple ways to the rest of the words stored in the lexicon, so that learning one lexical item has some effect on learning others (Meara & Wolter, 2004). Therefore, to develop full knowledge of a word it is necessary to build a rich and densely interrelated mental lexicon, which favors more rapid, comprehensive, and accurate networks between words (Cremer et al., 2010). However, examining the links among words is proving a very complex and challenging task.

It seems that, as in the L1, L2 learners develop their mental lexicon by adding and reorganizing the connections between words. Williams and Cheung (2011), in a semantic priming study on L1 Chinese learners of French, found that newly learned words did not simply adopt the L1 meanings. Rather, the new words automatically acquired their own semantic representations, which were associated with the contexts and meaning situations in which the words were learned. For example, when encountering a new word in a L2 (e.g., *écureuil* in French) and learning that its equivalent in the L1 is *squirrel*, one would expect that if the student knows that the word in the L1 (*squirrel*) is semantically associated with the word *nut*, then the new L2 word *écureuil* would also be. However, these authors found that the new word created its own associations based on the context in which it was learnt (e.g., a bushy tail based on a fairy tale character), and not on the meaning of the L1 word.

Therefore, L2 vocabulary learning is not seen as simply the integration of new knowledge into the existing L1 system, but as establishing connections between aspects of word knowledge through exposure to the word in varying contexts. That is, word knowledge develops from experiences and encounters with the language and connections between previous word knowledge and the new information, which will develop further with more and more varied exposures to a word (Perfetti, Wlotko, & Hart, 2005).

Table 16.1 Nation's (2013) framework of the dimensions involved in knowing a word

FORM	Spoken	[R]	What does the word sound like?
		[P]	How is the word pronounced?
	Written	[R]	What does the word look like?
		[P]	How is the word written and spelled?
	Word parts	[R]	What parts are recognizable in this word?
		[P]	What word parts are needed to express the meaning?
MEANING	Form and meaning	[R]	What meaning does this word form signal?
		[P]	What word form can be used to express this meaning?
	Concept and referents	[R]	What is included in the concept?
		[P]	What items can the concept refer to?
	Associations	[R]	What other words does this make us think of?
		[P]	What other words could we use instead of this one?
USE	Grammatical functions	[R]	In what patterns does the word occur?
		[P]	In what patterns must we use this word?
	Collocations	[R]	What words or types of words occur with this one?
		[P]	What words or types of words must we use with this one?
	Constraints on use	[R]	Where, when and how often would we expect to meet this word?
		[P]	Where, when, and how often can we use this word?

Note: [R] = receptive; [P] = productive.

Because word knowledge is acquired through multiple and varied language experiences (e.g., through both explicit instruction and incidental exposure: Schmitt, 2008), the acquisition of words is not a fixed process. Rather, word knowledge is a dynamic system that develops and changes over time, so that the acquisition of a word goes through different stages until all the word knowledge aspects needed to employ a word accurately in different situations (such as form–meaning mapping, collocational information, and word parts; see Table 16.1) are acquired (Fitzpatrick, 2012). This variable process makes it difficult to examine the links between words in the lexicon, and is one reason for the lack of a generally accepted theory of how the mental lexicon functions and vocabulary is acquired.

Key Concept

Mental lexicon: The mental dictionary where humans store the words they have some knowledge of. Those words are not stored individually, but appear to be highly organized and connected to each other in an intricate system. A rich and densely interrelated mental lexicon favors the development of depth of word knowledge.

Breadth and Depth of Word Knowledge

The terms *breadth* or *size* of vocabulary knowledge refer to the quantity of words a person has some knowledge of, and *depth* indicates the quality of that knowledge, that is, how well those words are known (Anderson & Freebody, 1981). It has been suggested that size and depth do not always grow in a parallel manner, because it is possible to learn a lot about a small number of words or a little about a large number of words (Schmitt,

2014). Nevertheless, the two dimensions are interrelated and contribute to one another (Li & Kirby, 2015; Qian, 2002; Schmitt, 2014; Tannenbaum, Torgesen, & Wagner, 2006). For example, the more words a learner knows (i.e., size), the more examples of word parts like prefixes and suffixes they will have in their mental lexicon, which in turn makes it easier for the learner to acquire the morphological aspects of vocabulary (i.e., depth). It is generally agreed that the development of depth of word knowledge is more problematic for learners and thus lags behind vocabulary size, regardless of the learners' proficiency level (see Schmitt, 2014). For example, Webb (2007) found that size (as represented by the form–meaning link) was generally learned earlier than depth (e.g., syntagmatic associations, paradigmatic associations, and word class). This gap is problematic because learners need to acquire depth of word knowledge to be able to use the words correctly, fluently, and appropriately in real situations.

Key Concepts

Word families: Lexical units that include all forms that share the same root plus all their inflectional and derivational affixes that (might) change the word's class (e.g., do, does, did, redo, undo, doable). The core meaning remains the same, although the form changes. This concept is used as a unit of vocabulary measurement, and gives lower numbers than research using individual words as vocabulary units.

Breadth and depth of word knowledge: In simple terms, breadth refers to how many words a person has some knowledge of (even if it is limited), and depth relates to the quality in which those words are known. Breadth has generally been conceptualized as knowledge of the form–meaning link of words (i.e., mapping a given L2 form to its meaning and/or an existing meaning to the appropriate L2 form). Depth, however, includes learning aspects such as the word class, collocations and grammatical functions, polysemous meanings, associations, and constraints on use.

Receptive and Productive Knowledge

Receptive knowledge refers to the learner being able to understand words encountered while reading or listening, and productive knowledge refers to using words in speaking or writing. Receptive mastery is typically reached before productive mastery, partly because productive mastery requires knowledge of more word knowledge aspects. Schmitt (2014) suggests that knowing the form–meaning link of a word might be enough for a receptive understanding of that word (although, of course, the more lexical aspects known, the better the comprehension is likely to be). In this situation, the user only needs to recall the meaning attached to the form that has been perceived, because all the other word knowledge aspects (e.g., word class, collocations, grammatical functions) are provided in the context. However, in order to produce a lexical item accurately and appropriately in a specific context, the user needs to know all (or as many as possible) of these aspects. That is, productive knowledge is more advanced than receptive knowledge (Read, 2000).

Different studies testing receptive and productive mastery of just the form–meaning link (e.g., Laufer & Paribakht, 1998; Tschirner, 2004) found that receptive mastery was higher than productive mastery (sometimes even five times higher: Nemati, 2010), and that, overall, when learners encounter higher frequency words, they are likely to both recognize and recall their form, whereas with low frequency words they can only recognize their form. Webb

(2007), using nonwords, studied the acquisition of receptive and productive mastery of various word knowledge aspects. He found that participants experienced gains in both receptive and productive knowledge, although the receptive knowledge of all the word knowledge aspects was always larger than the productive knowledge. This study was replicated by Chen and Truscott (2010) using real words, and the results were consistent. Moreover, a recent study by González-Fernández & Schmitt (under review), analysed the receptive and productive knowledge that L2 learners had of various word knowledge components, and found that the learners acquired receptive mastery in all the vocabulary knowledge aspects before productive mastery was achieved in any aspect. Therefore, the receptive knowledge of the different word knowledge aspects seems to be more robust and be acquired earlier than productive knowledge.

Learners Need a Large Vocabulary Size to Use a Language

One of the key issues in vocabulary teaching and learning is the amount of vocabulary L2 learners need to communicate. Research suggests that in order to communicate orally in basic, everyday informal situations, a vocabulary of between 2,000–3,000 word families in English is needed (if knowledge of roughly 95% of the vocabulary in the conversation is sufficient), or between 6,000 and 7,000 (assuming 98% coverage is needed) (Nation, 2006). There is not enough research to determine which of these coverage figures is sufficient, although van Zeeland and Schmitt (2013a) found that 95% was adequate for understanding informal narratives. The vocabulary requirements for reading are clearer, as by far the most research has been done in this area. Studies suggest that 8,000–9,000 word families (including proper nouns) provide 98% coverage, and are needed for L2 learners to read authentic texts (e.g., novels, newspapers) on a wide variety of topics in an independent manner. Knowledge of 4,000–5,000 families (with proper nouns) provides 95% coverage, which should enable initial engagement with these texts, albeit probably with the need for teacher support (Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006). It is very difficult to set size requirements for writing, as different writers are able to use the vocabulary they possess to better or worse effect (e.g., a person with a relatively smaller vocabulary may still be able to write convincingly if they use that vocabulary well).

These figures are for individual lexical items, and do not take into account lexical phrases or formulaic language. Consequently, they underestimate the true number of lexical items of various types that are necessary to communicate effectively. Thus, it is clear that, at least for reading and listening, learners need to acquire a large vocabulary to comprehend language efficiently.

Conceptualizing Depth: Aspects of Word Knowledge

Given its complexity, researchers have found it difficult to provide satisfactory descriptions of vocabulary depth. The most common framework is the components approach (Read, 2000), which describes the various components/aspects of word knowledge (e.g., form, meanings, word parts, collocations, and register) that make up the overall knowledge of a lexical item. This approach began as far back as 1942, when Cronbach recognized the multidimensional nature of word knowledge. In 1976, Jack Richards presented a list of eight assumptions involved in knowing a word, which was further developed by Nation in 1990. Nation's (2013) list is the most detailed and comprehensive conceptualization of word knowledge components to date (Table 16.1). In order to *fully* know a lexical item, the nine different aspects of word knowledge listed in Table 16.1 should be mastered, both receptively and productively.

Nation's conceptualization is currently widely used both by researchers examining the depth of knowledge construct, to select and describe the aspects of vocabulary knowledge to be assessed; and by teachers, because it can be applied to vocabulary learning in the classroom in a way that is relatively easy for both teachers and learners to understand.

Nation's framework is a maximal specification of lexical knowledge, and not even native speakers will have mastered all of these aspects for all the words they (partially) know. Therefore, L2 learners should not be expected to learn all their words in such depth. Nevertheless, gaining knowledge of even some of these aspects pushes students forward on the learning path (Schmitt, 2014).

Key Concept

Components approach: One approach for the description of word knowledge. It enumerates the different components of what it means to fully know a word, for example, form, meaning, grammatical characteristics, and constraints of use. The various components have been most fully specified by Paul Nation.

Vocabulary Acquisition Is Incremental in Nature

According to Schmitt (2010), vocabulary acquisition is incremental in many different ways. First, the various word knowledge aspects are not necessarily learned at the same rate. Rather, some aspects are learned before others and at different rates, although it is still very difficult to suggest an overall pattern, because few studies have examined the acquisition of multiple aspects concurrently. However, there have been some notable exceptions. Schmitt (1998) studied how different word knowledge aspects of 11 words (spelling, derivative information, associations, and polysemy) were acquired longitudinally. He found that as one of the aspects increased, so did the others, which suggests that the four word knowledge dimensions he explored were learned gradually and in a parallel manner. Webb (2007) used a battery of tests to examine the acquisition of five aspects of word knowledge productively and receptively (orthography, form–meaning link, syntax, grammatical functions and associations). Overall, he found parallel gains in all the different aspects although at different rates. Chen and Truscott (2010), following Webb's (2007) study, investigated the effect of repeated encounters on the acquisition of four word knowledge aspects (orthography, parts of speech, and associations both receptively and productively, and form–meaning link receptively), and found that increasing repetitions lead to better knowledge in all the different aspects, although the gains in knowledge varied depending on each aspect. In general, it has been found that the form–meaning link is one of the first aspects to be acquired in the process of vocabulary learning, and thus this aspect should be the initial target of L2 instruction (Schmitt, 2010). Aspects such as constraints of use or collocational knowledge have been found to be acquired later and require more time and many more exposures to develop.

Second, the development of each word knowledge aspect occurs incrementally. That is, these aspects are not learned in a dichotomous known/unknown fashion, but rather along a continuum, ranging from zero knowledge, to some partial knowledge to precise knowledge (Henriksen, 1999). For example, the knowledge of the spelling of a word can go from not knowing anything at all, to knowing just a few letters, then knowing some words with similar spelling, to finally acquiring the fully correct spelling.

Finally, the incremental nature of vocabulary acquisition is seen in the development of receptive and productive mastery. Research shows that learners' knowledge of vocabulary generally develops from receptive mastery to productive mastery (i.e., moving from ability to understand a word when listening or reading, to being able to produce it in speech or writing) (González-Fernández & Schmitt, under review; Nemati, 2010; Tschirner, 2004). Similarly, each aspect of word knowledge typically moves from receptive to productive mastery. The gradual transition from receptive to productive knowledge requires time and multiple exposures to a word, making this acquisition process incremental in nature.

However, Fitzpatrick (2012) warns that vocabulary acquisition does not always develop in a consistently upward trend. Rather, the process of vocabulary acquisition is somewhat unpredictable, and the knowledge of individual words and their word knowledge aspects will sometimes regress as well as move forward. For example, she found that the knowledge of the written form of a word tested at different points in time would move from generally correct spelling but with minor mistakes (**ture* instead of *true*) during the first testing period, to correct spelling (*true*) in subsequent periods, and then back to some minor spelling mistakes at a later point (**ture*), ending with correct spelling during the final test (*true*).

Formulaic Language Is Important

Vocabulary had traditionally been conceptualized as single words that were strung together by syntactical rules (Schmitt, 2010; Wray, 2002). However, corpus research has demonstrated that vocabulary consists not only of individual words, but also of large amounts of formulaic language (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Sinclair, 1991). *Formulaic language* is an overarching term used for various types of vocabulary (e.g., idioms, lexical bundles, collocations), which operate as multiword units. Formulaic language has been shown to be common in a range of languages, with estimates generally ranging from one-third to one-half of discourse for English (Conklin & Schmitt, 2012). It is so widespread because it carries out key communicative functions, such as in social interaction (*I understand, how nice*), for functional use (*how can I . . . I am sorry to hear that, I'd be happy to . . .*), and in organizing discourse (*on the other hand, in other words*). Formulaic language also has a key part in facilitating fluency, as it eases the processing and production of language, with less cognitive load for both the speaker and the interlocutor (Conklin & Schmitt, 2008; Pawley & Syder, 1983).

The diversity of formulaic language types makes it very difficult to define the concept and teach formulaic language, which is one reason why vocabulary instruction has traditionally focused mainly on teaching individual words, although with some exceptions, such as teaching basic functional phrases (e.g., for introductions, requests). But because formulaic language is central in language use, and because it has been found to pose problems for even advanced L2 learners (Levitzky-Aviad & Laufer, 2013; Nesselhauf, 2005; Paquot & Granger, 2012), it is important to incorporate formulaic language into vocabulary instruction.

Pedagogical Implications

Vocabulary is an essential aspect of language, but in many L2 classroom contexts, not much time is allocated to vocabulary teaching and learning. This lack of attention to vocabulary is a problem, because as Laufer and Nation (2011) point out, learning vocabulary entails the acquisition of thousands of items with many different aspects per item, and requires multiple encounters and considerable time. Laufer and Nation argue that vocabulary should thus be prioritized in the classroom. Because learners need to acquire a very large vocabulary to use a language well and because vocabulary is a complex construct of multidimensional

nature, language practitioners cannot assume that sufficient vocabulary can be acquired by simple exposure to grammatical or communicative activities. Rather, a comprehensive and principled vocabulary plan that involves explicit teaching in addition to exposure to large amounts of language needs to be implemented (Nation, 2013). Graves (2006) describes four main facets that must be included in any comprehensive approach to vocabulary instruction:

1. *Provide rich and varied language experiences.* Words have to be encountered by learners in listening, speaking, reading, and writing activities, in a variety of topics and genres. These activities require the involvement and independent work of students outside the classroom.
2. *Instruction of words.* New words need to be taught through direct instruction and explicit methods, using clear explanations and simple definitions first, and then providing more information when the words are recycled and encountered again in varied contexts. Different teaching methods should be used depending on the characteristics of the words to be learned, and the stage of learning.
3. *Teaching strategies for autonomous vocabulary learning.* Learners need to be taught how to infer words from context or morpheme clues, to use dictionaries, and to connect the knowledge of new words with previously known words.
4. *Foster the active engagement of students in vocabulary learning.* Use activities that promote the interest and involvement of learners, and motivate them to learn more about words. Based on a Structural Equation Model, Tseng and Schmitt (2008) demonstrate that learners' motivation is involved in all stages of vocabulary learning, and thus is crucial to a beneficial vocabulary learning process.

Graves's approach involves both intentional and incidental learning, which will be the next issues addressed in this section.

Incidental Learning of Vocabulary

Incidental vocabulary learning refers to the process of acquiring vocabulary knowledge when the specific lexical item being learned is not the main focus of either the teaching or learning activity (Ender, 2016). The learners' purpose is enjoying the task or understanding a specific message, but in this process they acquire some words without making a conscious effort (Ellis, 1994; Hulstijn, 2001). It is clear that substantial vocabulary can accrue incidentally through reading or other activities (Chen & Truscott, 2010; Ender, 2016; Gass, 1999; Hulstijn, 2013), although the uptake rate is generally slower and more uneven than with intentional learning.

A key issue is the number of exposures necessary to learn vocabulary incidentally from context. Study results vary widely depending on what aspects of vocabulary are measured, but as a rule of thumb, 8–10 exposures from reading seem sufficient for learners to be able to answer form–meaning multiple-choice vocabulary items correctly in subsequent tests (Schmitt, 2008), or to read new words as quickly and accurately as previously known words as evidenced by eye-tracking (Pellicer-Sánchez, 2016). However, some researchers have suggested that as few as three encounters are enough for learners to acquire the meaning of a target word if the reading is important and interesting for the students (Reynolds, Wu, Liu, Kuo, & Yen, 2015). There is little research on incidental learning from listening, but one study suggests that for listening to be a valuable source for vocabulary learning (specifically of form and meaning), considerably more than 15 exposures may be needed (van

Zeeland & Schmitt, 2013b). This finding implies that durable and meaningful incidental vocabulary acquisition from listening seems to require more exposures than from reading.

In order to ensure this repeated contact with words, teachers need to find ways to increase students' L2 exposure inside and outside the classroom, and one of the most common ways of doing this is by extensive reading, which is considered a very positive way of increasing and improving learners' L2 vocabulary (Uden, Schmitt, & Schmitt, 2014). However, even in extensive reading, frequency plays a big role in learnability; while high-frequency words appear often enough to have a good chance of being acquired, lower frequency words do not. Cobb (2007), in a corpus-based study, found that words beyond the most frequent 2,000 level will be met rarely, if at all, in the period of a year even with relatively large amounts of reading exposure. Thus, to learn new vocabulary, incidental learning is not enough: explicit instruction that may lead to intentional learning is also required.

Teaching Tips

- Include an extensive reading (e.g., graded readers) component to your language curriculum to maximize the amount of incidental vocabulary learning.
- Vocabulary knowledge aspects such as constraints of use or collocations have been found to require many more exposures than aspects such as form and meaning. Thus, such aspects are good candidates for incidental learning from massive exposure.

Intentional Vocabulary Learning

Intentional vocabulary learning refers to the deliberate attempt to learn new words (Hulstijn, 2005), and it involves acquiring new vocabulary through direct instruction and the use of personalized vocabulary learning strategies. Examples of learning activities include word flashcards, multiple-choice activities, matching words, and fill-in-the-blank exercises. Research shows that deliberate, intentional vocabulary teaching and learning can increase vocabulary knowledge quickly and effectively (e.g., Webb, 2007). Intentional learning has also been found to lead to better results than incidental learning (Cobb, 2007; Horst, Cobb, & Nicolae, 2005; Joyce, 2015). Laufer and Rozovski-Roitblat (2011) investigated how intentional and incidental activities influenced the acquisition of new words. They found that intentional activities (i.e., practicing decontextualized vocabulary by matching written word forms with their definitions, synonyms and antonyms; selecting the correct meaning from various options; and writing target words in sentences) were more effective than incidental tasks for vocabulary learning regarding recognition of meaning and form; moreover, such activities lead to long-term retention.

There is an almost unlimited number of potential vocabulary learning activities, and we do not have a clear idea of their relative effectiveness. However, virtually any activity that leads to more exposure, attention, manipulation, or time spent on lexical items seems to facilitate learning (Schmitt, 2008). Even vocabulary testing and other activities that some would consider old-fashioned and out of date can be effective. Bilingual word lists or word cards, for example, have been found to be effective in the acquisition and retention of newly learned words, both productively and receptively (Yamamoto, 2014). One of the underresearched, but promising, vocabulary learning activities involves meaning-focused

output, in which learners are encouraged to use vocabulary in new contexts. Meaning-focused output is beneficial to vocabulary acquisition in three ways: it encourages the use of new vocabulary and the negotiation of the meaning of unknown vocabulary, and strengthens learners' knowledge of partially known items by using them in language production (Nation & Meara, 2002).

Intentional, word-focused vocabulary acquisition is effective in increasing learners' vocabulary size and depth (Yamamoto, 2014), but intentional activities need to be combined with incidental, contextualized, message-focused activities, because the latter help consolidate the previous knowledge (often initially learned through direct study), as well as develop further depth of word knowledge (Joyce, 2015; Laufer & Nation, 2011). Also, some word knowledge aspects are better learned through explicit study (e.g., form–meaning link, word parts), while others require exposure to many instances in a variety of contexts (e.g., collocations, register). Thus, the current best-practice approach to vocabulary instruction combines both intentional and incidental learning (Nation, 2013).

Teaching Tip

Learners can learn much vocabulary on their own. Look at your materials in advance and determine the words your students are unlikely to know. Fix these to a word list and have your students study them before the class. Then when you use the words in readings and examples, your students will be better able to understand them in their contextualized settings.

Multiple Encounters With a Word Are Necessary

In vocabulary instruction, the form–meaning link is considered the most important component, because it is the first one to be developed and is the minimum aspect needed for communication. Thus, the central focus of vocabulary teaching in the first instance must be the form–meaning link. However, it must not be forgotten that knowing vocabulary involves more than just being able to make that link. If L2 learners are to be able to use the target language appropriately, vocabulary instruction must also subsequently focus on enhancing as many aspects of word knowledge as possible, which requires many and varied encounters with a word.

Recycling of a target word has been found to improve knowledge of the various aspects of word knowledge for that word, both productively and receptively. Webb (2007) examined how Japanese EFL students acquired nonwords from different exposures (1, 3, 7 and 10). He measured five aspects of word knowledge (orthography, form–meaning link, syntax, grammatical functions and associations). Overall, he found that the more exposures to a word, the better the gains in all the different aspects. Chen and Truscott (2010) studied the effect of repeated encounters (1, 3 and 7) on the acquisition of four word knowledge aspects (orthography, parts of speech, associations, and form–meaning link). Similarly, they found that increasing repetitions led to better knowledge in all the different aspects, although the effect of repetition varied depending on the aspect.

From even one encounter with a word, learners can pick up initial information about the form–meaning link, and thus increase their vocabulary size (Webb, 2007). However, more repetitions are needed for that knowledge to settle, and with those repetitions other aspects of word knowledge develop. Pellicer Sánchez and Schmitt (2010) found that at 10

or more encounters with a word, substantial gains occurred in word form recognition, word class recall, and meaning recognition and recall. Thus, as vocabulary size increases, so does vocabulary depth (although usually with a time lag), and exposure to the L2 not only allows the learning of new words, but also reinforces and increases the depth of knowledge of other words (Qian, 1999).

Multiple encounters with a word also helps student's knowledge develop from receptive to productive mastery; furthermore, language practitioners need to understand that their learners' receptive/productive profile is likely to vary considerably according to the number of exposures to those words (Chen & Truscott, 2010; Webb, 2007).

Overall, recycling is fundamental to effective vocabulary instruction, and teachers should provide opportunities/activities that allow students to encounter a word repeatedly and in varied contexts, to both consolidate and enhance their understanding of it.

Teaching Tip

Textbooks usually do not recycle words to any great extent. The creation of supplementary materials (e.g., word games, speaking activities with a target word list) focusing on already-taught words will aid in their retention and elaboration.

Selection of Words

In order to decide what vocabulary to focus on in language teaching, there are some principles that L2 teachers can follow.

Because the vocabulary of L2 learners is limited, teachers should teach those words which are as useful as possible for the learners. This criteria means that the selection of words for instruction should be based on: frequent words that students will encounter often, generalizable words that are useful for various purposes, words that are less frequent but attend to the students' personal needs, and learnability of words (i.e., words considered easier or more difficult for students). For example, cognates (words similar in form and meaning between two languages) and concrete words seem to be easier for students than false cognates (words similar in form but different in meaning) or abstract words (Graves, 2006; Laufer & Nation, 2011).

Frequency counts are considered one of the best ways of selecting the vocabulary that will be most valuable for learning. From a cost-benefit perspective (Nation, 2013), high-frequency words give a better return for learning than low-frequency words in any language, and therefore are the ones teachers should focus on. This strategy seems to be particularly true for the earliest learning stages, because it has been suggested that the most frequent 3,000 words are essential in English, and thus students benefit greatly from knowing them. High-frequency vocabulary allows learners to understand most of the words they are exposed to, because they account for around 90% of written and spoken English (Nation, 2006).

However, learning the first 3,000 high-frequency word families in English is only the beginning, and teachers and learners need to focus on other words beyond this level. Schmitt and Schmitt (2014) argue that learners also need large amounts of mid-frequency vocabulary (3,000–9,000) to function well in English. Beyond this, Nation (2013) believes that low-frequency words (9,000+) occur too rarely to warrant the cost of teaching them. For these words, teachers should focus on instructing and encouraging students to use learning

strategies (e.g., guessing from context or morphology, mnemonics) so that they can acquire these low-frequency words on their own.

Nevertheless, frequency is not the only criterion for the selection of words in language teaching. Teachers also need to focus on the particular needs of learners (e.g., spoken vocabulary) and words that are useful in specific contexts (e.g., technical vocabulary). In fact, Schmitt (2010) suggests that mastering technical vocabulary is the logical next step after a person knows the first 5,000 word families. Teachers can also consider the learnability of words. Cognateness (having similar form/meaning in the L1) is one of the best predictors of learning (Willis & Ohashi, 2012), and so cognates can be good candidates for attention. Teaching cognates could either entail explicit instruction or awareness-raising, which could facilitate learners recognizing and understanding L2 cognate items (Bahns, 1993).

Teachers can use word lists to guide their vocabulary selection, as long as the lists match the teachers' specific pedagogical purposes. Some useful lists available include two New General Service Lists (NGSL), for general high-frequency vocabulary (Brezina & Gablasova, 2015; Browne, 2013), the Academic Vocabulary List for academic vocabulary (Gardner & Davies, 2014), the PHRASE (PHRASal Expressions) List for frequent formulaic sequences (Martinez & Schmitt, 2012), and the PHaVE (PHRASal VERb) List for phrasal verbs (Garnier & Schmitt, 2015).

Teaching Tip

There are some very useful, freely available vocabulary lists that can support vocabulary teaching, such as the PHaVE, PHRASE, NGSL, and New Academic World List (NAWL) lists.

Nation's Four Strands of Vocabulary Instruction

To sum up, this review has argued that vocabulary is a complex construct with different aspects and characteristics that require various approaches and techniques to be acquired. Therefore, a good vocabulary instruction program should take into account and balance all these different methods to lead to a comprehensive vocabulary experience. With this view, Nation (2007) suggests a four-strand approach to a well-balanced vocabulary course.

1. *Learning from comprehensible, meaning-focused input.* This refers to learning vocabulary through reading and listening activities, where the main focus is on understanding, gaining information, or enjoying the activity. This strand is directly connected to incidental learning and the receptive use of language, where learners acquire some knowledge of new words through context. Some common activities include watching TV or films, extensive reading, teacher's input in the classroom and role play conversations. However, in order for this approach to be effective, learners need to know (at least to the form–meaning level of mastery) most of the words used in the readings or listening activities (around 95–98%) and need to be interested and motivated to do the activity.
2. *Learning from meaning-focused output.* In this strand, learning occurs through speaking and writing, where the main focus of attention is not accuracy or correction, but using the language for communication to convey a specific message. Some typical activities would be conversations, writing a letter, telling or writing a story, or giving a talk. The use of this message-focused output provides learners with different learning opportuni-

ties than those provided by input. For example, output activities can help learners notice gaps in their productive vocabulary knowledge (i.e., being conscious of their lack of productive mastery of certain words), as well as take the risk of using words they are not completely sure about, which will confirm or change what they previously knew about the use of those words (Swain, 1995, 2005). Speaking and writing activities help learners focus on the productive aspect of words they know receptively.

3. *Learning from language-focused or form-focused instruction.* This strand involves the direct teaching and learning of vocabulary and its different aspects, such as spelling, pronunciation, grammatical features, or discourse features. Some typical activities are matching or fill-in-the-blank tasks (Laufer & Rozovski-Roitblat, 2011), using word cards or word lists to learn vocabulary (Elgort, 2011), practicing pronunciation (de la Fuente, 2002), translation (Joyce, 2015; Laufer & Girsai, 2008), and explicitly using glosses (Hulstijn, Hollander, & Greidanus, 1996) or dictionaries (Scholfield, 1997) to learn new words.
4. *Fluency development.* This strand is connected with the four skills (listening, reading, writing, and speaking), where the focus is to receive and convey messages, without worrying about accuracy. It entails having learners use their previously (but partially) learned vocabulary in timed activities in order to develop and enhance fluency of use, that is, the ability to utilize vocabulary in real-time use. Some common activities include skimming and scanning, speed reading, and timed writing. In this strand, all the vocabulary the learners are using or exposed to must be known, because the focus is using the language they already (partially) know more fluently, not learning new words. In this sense, fluency of use could be considered part of depth of word knowledge.

Nation (2007) suggests that, in a well-designed language course, these four strands should be given equal amounts of time, about 25% of the course time. This way, an appropriate balance of learning opportunities is provided, covering both receptive and productive skills. However, this ratio will depend on the teaching context. For example, beginners are likely to benefit from a greater proportion of language- and form-focused instruction, while intermediate/advanced learners should have a larger vocabulary size, which enables learning from meaning-focused input and output activities.

Teaching Tip

Use of Nation's four strands can ensure that learners receive a well-rounded range of input and output opportunities to learn and use vocabulary.

Future Research

Despite the large amount of research conducted in vocabulary acquisition during the past 30 years, there are still many issues for which little or nothing is known, and thus, require further research.

Theoretical Issues

Due to the complexity of the vocabulary construct, research has not focused enough on the networks between words and the links between different aspects of knowledge of an individual word. As a consequence, there is a lack of a generally accepted theory of vocabulary

acquisition. Research should focus on assessing the various aspects of word knowledge concurrently with a battery of tests in order to better understand their relationships and development. An example of this approach would be developing a series of both productive and receptive tests specifically designed to assess the knowledge of various aspects of word knowledge, and submit the results to statistical analyses that show causal connections between the different components (i.e., Structural Equation Modelling) (González-Fernández & Schmitt, under review). Knowledge of these connections between aspects would shed light on the overall process and nature of vocabulary acquisition and would allow the development of a multidimensional theory of L2 vocabulary learning, and the principled and systematic teaching of vocabulary.

Assessment Issues

In order to examine the acquisition of vocabulary, it is necessary to develop measurement instruments targeted at the different aspects of word knowledge at different levels of sensitivity. Some aspects, like the form–meaning link (especially in the written mode), have been well-researched. However, there are other components that have hardly been studied (e.g., constraints on use), and thus no commonly accepted measures have been developed. Therefore, research needs to focus more on the creation of standardized tests to assess various aspects of depth of vocabulary knowledge.

Instruction Issues

Regarding vocabulary instruction, there are certain issues that need to be addressed. It is now clear that some teaching activities are more effective than others for vocabulary instruction. For example, reading activities combined with learning from cards, matching words, multiple-choice activities, and writing unrelated sentences yield better learning outcomes than using dictionaries while reading (Laufer & Rozovski-Roitblat, 2011). Similarly, oral tasks where learners attend to unknown words by asking for clarification usually yield better retention of the target words than when learners do not draw attention to the words (de la Fuente, 2002). However, this effectiveness depends on many factors such as level of engagement, amount of time spent, and proficiency level. Thus, it is still unclear which specific characteristics of explicit teaching activities make them more effective. Research could be usefully done to identify the most important features of effective activities and how they relate to the teaching of the different aspects of depth of word knowledge.

There is also little understanding about how the different types of formulaic language are learned and retrieved from memory. This insufficient knowledge has resulted in a lack of a principled approach to teaching formulaic language. Because formulaic language is a key component of language use, further research on how to best teach it is necessary.

Conclusion

A large vocabulary is necessary to use an L2 well. Given the number of lexical items that need to be learned (both words and formulaic language), only a principled approach to teaching these items will be successful. First, there should be a sensible selection of the vocabulary to be taught, based on the learnability and frequency of words, but also learners' needs. Second, once the vocabulary for instruction is decided, teachers need to draw attention to the acquisition of not only size of vocabulary, but also depth. This involves learners

gaining a range of word knowledge aspects (e.g., collocations, word parts, grammar) to receptive and productive mastery. Finally, this approach needs to provide a variety of learning opportunities by combining and balancing the best of explicit teaching and the benefits of incidental learning from recycling vocabulary through varied and very large amounts of language exposure.

References

- Alderson, J. C. (2007). Judging the frequency of English words. *Applied Linguistics*, 28(3), 383–409.
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.
- Bahns, J. (1993). Lexical collocations: A contrastive view. *English Language Teaching Journal*, 47(1), 56–63.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the New General Service List. *Applied Linguistics*, 36(1), 1–22.
- Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning. *Language Teacher*, 37(4), 13–16.
- Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, 31(5), 693–713.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology*, 11(3), 38–63.
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than non-formulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72–89.
- Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, 32, 45–61.
- Cremer, M., Dingshoff, D., Beer, M., & Schoonen, R. (2010). Do word associations assess words knowledge? A comparison of L1 and L2, child and adult word associations. *International Journal of Bilingualism*, 15(2), 187–204.
- Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *Journal of Educational Research*, 36, 206–217.
- de la Fuente, M. J. (2002). Negotiation and oral acquisition of L2 vocabulary (The roles of input and output in receptive and productive acquisition of words). *Studies in Second Language Acquisition*, 24(1), 81–112.
- Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, 61, 367–413.
- Ellis, N. (1994). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211–282). London: Academic Press.
- Ender, A. (2016). Implicit and explicit cognitive processes in incidental vocabulary acquisition. *Applied Linguistics*, 37, 536–560.
- Fitzpatrick, T. (2012). Tracking the changes: Vocabulary acquisition in the study abroad context. *Language Learning Journal*, 40(1), 81–98.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305–327.
- Garnier, M., & Schmitt, N. (2015). The PHaVE List: A pedagogical list of phrasal verbs and their most frequent meaning senses. *Language Teaching Research*, 19, 645–666.
- Gass, S. (1999). Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21(2), 319–333.
- González-Fernández, B., & Schmitt, N. (under review). Order of acquisition and relationships among word knowledge components.

- Graves, M.F. (2006). *The vocabulary book: Learning and instruction*. New York: Teachers College Press.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21, 303–317.
- Horst, M., Cobb, T., & Nicolae, H. (2005). Expanding academic vocabulary with an interactive on-line database. *Language Learning and Technology*, 9(2), 90–110.
- Hulstijn, J.H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge: Cambridge University Press.
- Hulstijn, J.H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition*, 27(2), 129–140.
- Hulstijn, J.H. (2013). Incidental learning in second language acquisition. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2632–2637). Oxford: Wiley-Blackwell.
- Hulstijn, J.H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal*, 80(3), 327–339.
- Joyce, P. (2015). L2 vocabulary learning and testing: The use of L1 translation versus L2 definition. *Language Learning Journal*, 43, 1–12.
- Larsen-Freeman, D. (2000). *Techniques and principles in language teaching*. Oxford: Oxford University Press.
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694–716.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399–436.
- Laufer, B., & Nation, I.S.P. (2011). Vocabulary. In S.M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163–176). London: Routledge.
- Laufer, B., & Paribakht, T.S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, 365–391.
- Laufer, B., & Ravenhorst-Kalovski, G.C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30.
- Laufer, B., & Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research*, 15, 391–411.
- Levenston, E.A. (1979). Second language acquisition: Issues and problems. *Interlanguage Studies Bulletin*, 4(2), 147–60.
- Levitzky-Aviad, T., & Laufer, B. (2013). Lexical properties in the writing of foreign language learners over eight years of study: Single words and collocations. *EUROSLA Monographs Series*, 2, 127–148.
- Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, 36, 611–634.
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299–320.
- Meara, P.M. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching*, 13(3–4), 221–246.
- Meara, P.M., & Wolter, B. (2004). V LINKS: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English speaking world 4* (pp. 85–96). Copenhagen: University of Copenhagen/ Museum Tusculanum Press.
- Nation, I.S.P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82.
- Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 1–12.
- Nation, I.S.P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.

- Nation, I.S.P., & Meara, P. (2002). Vocabulary. In Norbert Schmitt (Ed.), *An introduction to applied linguistics* (pp. 35–54). London: Arnold.
- Nemati, A. (2010). Active and passive vocabulary knowledge: The effect of years of instruction. *Asian EFL Journal*, 12, 30–46.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam/Philadelphia: John Benjamins.
- Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130–149.
- Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191–226). London: Longman.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, 38, 97–130.
- Pellicer Sánchez, A. M., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things Fall apart? *Reading in a Foreign Language*, 22, 31–55.
- Perfetti, C.A., Wlotko, E. W., & Hart, L.A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning Memory and Cognition*, 31, 1281–1292.
- Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56, 282–308.
- Qian, D.D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52, 513–536.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Reynolds, B. L., Wu, W. H., Liu, H. W., Kuo, S. Y., & Yen, C. H. (2015). Towards a model of advanced learners' vocabulary acquisition: An investigation of L2 vocabulary acquisition and retention by Taiwanese English majors. *Applied Linguistics Review*, 6(1), 121–144.
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10(1), 77–89.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281–317.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Houndmills, Basingstoke: Palgrave Macmillan.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913–951.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.
- Scholfield, P. (1997). Vocabulary reference works in foreign language learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 279–302). Cambridge: Cambridge University Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Lawrence Erlbaum Associates.
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10(4), 381–398.
- Tschirner, E. (2004). Breadth of vocabulary and advanced English study: An empirical investigation. *Electronic Journal of Foreign Language Teaching*, 1, 27–39.
- Tseng, W.-T., & Schmitt, N. (2008). Towards a self-regulating model of vocabulary learning: A structural equation modeling approach. *Language Learning*, 58(2), 357–400.

Beatriz González-Fernández and Norbert Schmitt

- Uden, J., Schmitt, D., & Schmitt, N. (2014). Jumping from the highest graded readers to ungraded novels: Four case studies. *Reading in a Foreign Language*, 26(1), 1–28.
- van Zeeland, H., & Schmitt, N. (2013a). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34, 457–479.
- van Zeeland, H., & Schmitt, N. (2013b). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, 41, 609–624.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46–65.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Williams, J. N., & Cheung, A. (2011). Using priming to explore early word learning. In P. Trofimovich & K. McDonough (Eds.), *Applying priming methods to L2 learning, teaching and research* (pp. 73–104). Amsterdam: John Benjamins.
- Willis, M., & Ohashi, Y. (2012). A model of L2 vocabulary learning and retention. *Language Learning Journal*, 40(1), 125–137.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Yamamoto, Y. (2014). Multidimensional vocabulary acquisition through deliberate vocabulary list learning. *System*, 42, 232–243.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 5–19). Cambridge: Cambridge University Press.

Written Language Learning

Charlene Polio and Jongbong Lee

Background

A discussion of the acquisition of L2 written language in instructed contexts requires a description of the scope of the research. First, much L2 writing research focuses on discourse-level issues and may examine, for example, how students develop genre knowledge (e.g., Tardy, 2012) or how voice and identity are realized in L2 writing (see Matsuda, 2015, for a review). Although concepts such as genre or voice are often discussed in terms of lexical and morphosyntactic features, those features are not the focus of SLA research. Therefore, in this chapter, we limit our review to research that examines L2 knowledge as represented in written production. Second, it is somewhat difficult to pinpoint what research falls under the realm of *instructed* SLA (ISLA). On one hand, virtually all L2 writing research includes participants who have been instructed at some point. On the other hand, much writing research focuses on *previously* instructed learners who write in contexts outside of language classes and are currently not enrolled in L2 classes (e.g., Li & Schmitt, 2009; Lillis & Curry, 2010). In this chapter, we limit our review to studies that occur in classrooms or are conducted with currently instructed learners and studies that include “the manipulation of the mechanisms of learning and/or the conditions under which they occur” (per Loewen’s definition of instruction, 2015, p. 3).

Written SLA research has a comparatively shorter history than spoken SLA. Krashen (1982), for example, highlighted the role of comprehensible input for oral skills; writing was a way for learners to monitor their language and to apply their knowledge of rules, but did not have a place in his theory. Long (1996), criticizing the singular role of comprehensible input, focused on oral interaction as a necessary component for acquisition. He highlighted the role of immediate feedback in the form of recasts and negotiation, feedback generally not possible during the individual production of written texts.

Swain (1985) also took Krashen’s focus on oral language as a starting point in her influential work. She argued that comprehensible *output*, in addition to input, is necessary for learning; learners acquire language by trying to make themselves understood. With this new role for output, written language production became more relevant. Swain and Lapkin

Charlene Polio and Jongbong Lee

(2002) stated, “Our program of research has focused on the roles of output (speaking and writing) in second language learning” (p. 285), thereby specifically including written language in the discussion of how languages were learned. Swain and Lapkin (1995) examined anglophone adolescents thinking aloud as they wrote in French. The students were found to notice and solve some language problems as they wrote. Swain and Lapkin stated, “What goes on between the first output and the second, we are suggesting, is part of the process of second language learning” (p. 386). Earlier, Cumming (1990) had noted the importance of writing as a way to draw learners’ attention to linguistic form as they created meaning. He said:

Composition writing elicits an attention to form–meaning relations that may prompt learners to refine their linguistic expression—and hence their control over their linguistic knowledge—so that it is more accurately representative of their thoughts and of standard usage.

Cumming, 1990, p. 483

Swain and Lapkin’s work confirmed this view and from then on, writing was seen as a way for L2 learners to focus on language forms because it afforded learners an opportunity to pause, monitor, and repair their language, processes that could be considered markers of dysfluency in speaking.

These discussions focused more on language production than on instruction, but they allowed for speculation on the role of writing in SLA. The idea that writing might facilitate SLA seemed to validate the teaching of writing even in contexts in which students have undefined real-life writing goals. Manchón (2011), and others, have identified the *writing-to-learn-language* approach as an important way to teach language. Ortega (2011) stated that writing-to-learn-language “seeks to carve out a substantive and valued role in L2 classrooms, elevating it from a convenient way to practice grammar and vocabulary to a site for language development” (p. 240). As we show in this chapter, both written production and writing instruction may facilitate L2 development even if the links are not always direct.

Key Concept

Writing-to-learn-language: This refers to either a classroom activity or a general approach to using writing as way to teach or to have students practice language. The activities generally do not reflect real-world writing tasks (e.g., writing a cover letter) and instead may include an activity such as having students describe a picture using a set of vocabulary.

Current Issues

We begin with a discussion of the role of writing in SLA followed by an overview of the debate regarding written corrective feedback. Although research on corrective feedback has been discussed extensively in many publications (for a book length review, see Bitchener & Ferris, 2012), it remains a major concern. We then discuss how different writing tasks or prompts may affect production and learning, and how language develops in specific instructional

contexts. This section discusses the scope of the issues, and in the next section, we detail empirical studies examining the issues.

The Role of Writing in SLA

It is accepted that oral and literacy skills are related in some way (Belcher & Hirvela, 2008). For example, Bigelow, Delmas, Hansen, and Tarone (2006) found that nonliterate students had more difficulty recalling feedback from oral recasts and suggested that first language (L1) literacy skills may help learners process second languages. With regard to writing, Harklau (2002) argued that “writing should play a more prominent role in classroom-based studies of second language acquisition” (p. 329). She found in her classroom-based research that ESL students learned from written input, as evidenced by their ability to spell words. She did not document how learning to write transferred to oral skills, but she cogently argued that the writing was a site of language acquisition: students wrote more than they spoke, they received more feedback in writing, and it allowed them to edit and monitor their language.

Weissberg (2000) argued that writing might be the preferred modality for the use of new grammatical forms. He followed five adult ESL learners as they completed a variety of oral and written tasks. It appeared that more, but not all, structures appeared first in writing and that learners were more accurate in writing. In Weissberg (2006), he discussed moving from speaking to writing in the classroom suggesting that the relationship may be bidirectional. This speaking-to-writing direction has long been discussed in the L2 (and L1) writing literature; oral discussion has been viewed as an effective prewriting activity (see Shi, 1998, for a review). However, there is little discussion in the pedagogical literature of writing as a prespeaking activity despite some empirical evidence provided in the next section that suggests writing before speaking may be helpful.

Teaching Tip

- Have students do a related writing activity before they do an in-class oral activity.

The most direct discussion of the learning potential of writing is Williams (2012). Williams drew on Housen and Pierrard’s (2005) model of L2 development to explain how writing can facilitate acquisition at various stages of the learning process and how writing might be superior to oral production in this respect. Specifically, she drew on learners’ stages of knowledge internalization, restructuring, and consolidation to show how they apply to written production. She argued that because writing is permanent (i.e., there is visual record) and slower than speaking, there is “more learner control over attentional resources as well as more need and opportunity to attend to language both during and after production” (p. 323). Williams also drew on Laufer and Hulstijn’s (2001) involvement load hypothesis supporting the finding that learners retained vocabulary better after writing than reading. She also explained that writing activities could help learners create new knowledge during the internalization and restructuring phases of SLA. Many studies (e.g., Brooks & Swain, 2009; Gutiérrez, 2008; and others reviewed in Storch, 2011) have indeed shown that as learners write together, they co-construct L2 knowledge that then appears in their writing.

Key Concept

Involvement load hypothesis: This hypothesis states that the more learners are involved with vocabulary, the better they will retain it. Involvement includes factors such as having to search for a word's meaning and having to use a word to convey meaning. It has been argued that having to use new words in an essay creates a great amount of involvement.

Teaching Tips

- Use writing activities even if students are generally more interested in developing their speaking skills.
- Have students work together on some writing activities.

Corrective Feedback

The relationship between written corrective feedback and L2 development in writing classrooms is a major issue. On one hand, the role of error correction has recently been downplayed in some of the teaching methods books in favor of more global or content-related feedback (e.g., Williams, 2005; Weigle, 2014), but written corrective feedback has been widely researched as evidenced by several meta-analyses on the topic, discussed in the next section.

Key Concept

Written corrective feedback: This is feedback on language as opposed to global feedback, which may focus on issues such as content or organization. Written corrective feedback may include direct correction of an error, coding, or underlining an error.

Written corrective feedback came under scrutiny when Truscott (1996) argued that it was not effective and should be abandoned. He argued that some empirical studies (e.g., Robb, Ross, & Shortreed, 1986; Semke, 1984) had shown that such feedback was not helpful, and that studies claiming effectiveness were flawed (e.g., Fathman & Whalley, 1990; Lalande, 1982). Indeed, the lack of methodological rigor in many of the studies suggests that it was difficult to draw any firm conclusions from the research (Bruton, 2009; Ferris, 1999; Guénette, 2007; Liu & Brown, 2015; Polio, 1997, 2012a; Xu, 2009). Better designs including the use of reliable measures, comparable control groups, assessment of long-term effects, explicit definitions of feedback, and attention to the effect of feedback on all aspects of writing, not only accuracy, are reflected in the studies discussed in the next section.

Truscott (1996) also argued that given what was known about SLA, error correction was not expected to be effective. Drawing on research from developmental sequences (e.g., Dulay & Burt, 1973, 1974; Pienemann, 1984, 1989), Truscott claimed that corrective feedback could not alter the natural process of acquisition. Truscott also said that corrective

feedback could lead only to what he called *pseudolearning* (presumably explicit knowledge). Polio (2012b) addressed these concerns by examining how different theories of SLA might view corrective feedback through different theories of SLA. She concluded that while some theories had little to say about the role of written error correction in language learning, some suggested that it could be effective under certain conditions. One example, skill acquisition theory, best represented by the work of DeKeyser (2007, 2015), suggests a role for feedback and explicit knowledge in SLA, and this theory was applied to an empirical study of feedback discussed in the next section.

The Effect of Different Tasks on Written Language Production

If writing facilitates SLA, we should understand how writing prompts or tasks affect written language production. In other words, if students use more complex language while performing certain tasks, it might be best to use those tasks in the classroom to push development. Many empirical studies on written tasks and production (e.g., Frear & Bitchener, 2015; Kuiken & Vedder, 2008, both discussed in the next section) were influenced by theories originally associated with spoken SLA. Such research has revolved around two conflicting theories, the cognition hypothesis (Robinson, 2001) and the limited attention capacity model (Skehan, 1998). To simplify, the former suggests that if learners are provided with a complex task, such as one that has more reasoning demands or more elements to keep track of, they will produce more accurate and complex language. The latter hypothesizes that a more complex task will divert learners' attention away from producing more complex language. Both, however, predict that additional planning time, which is also a task condition, increases linguistic complexity, accuracy, and fluency. One implication of task research is that instructors might be able to get students to produce more accurate and complex language with certain task types.

The application of a theory intended for the acquisition of oral language to written, however, may be problematic. Jackson and Suethanapornkul (2013) voiced concerns about the application of the cognition hypothesis to writing tasks. In their meta-analysis of studies testing the cognition hypothesis, they excluded studies of task complexity in the written mode because the planning variable was too difficult to control. In addition, how task and genre variables in writing interact is not clear. Yoon and Polio (2017) and others (e.g., Lu, 2011) have found differences in linguistic complexity across genres. Specifically, students use more complex language in argumentative essays than in narratives. One interpretation is that the argumentative essays are more complex because of additional reasoning demands, but another is that one simply needs more complex language because of the communicative demands of the genre, as explained by Biber and Conrad (2009).

Some of the research on task differences is situated not in the task or SLA literature but in the assessment literature because test writers want to use prompts that represent the kind of language elicited on a comparable real-life writing task or because they do not want to use prompts that unfairly give an advantage to some students. He and Shi (2012), for example, studied how prompts related to general knowledge for university students (i.e., factors influencing their major) versus specific knowledge (i.e., their interest in federal politics) affected their writing. They investigated linguistic features, namely, accuracy, the use of academic words, and fluency, as measured by essay length. They found students' scores were higher on all linguistic features when responding to the prompt related to general knowledge as opposed to a topic with which they were less familiar. The authors never framed nor explained the study in relation to SLA principles and instead focused on issues of

Charlene Polio and Jongbong Lee

validity and fairness in testing, but we can turn to research on task complexity for a possible explanation. Specifically, students likely had more time to plan aspects of their essays other than content in the familiar condition thus leading to differences in the produced language.

Taken together, the research on task complexity, genre, and assessment all show that learners' language varies in relation to what they are asked to write, although we cannot be completely sure why this variation occurs. Furthermore, we do not know much about the long-term effects of these tasks or if any learning can be transferred to oral language. In the next section, we review some of these studies and then later, explain the instructional implications of language differences across writing tasks.

Key Concept

Prompts, tasks, and genres: These terms all refer to ways to elicit written language for research or testing purposes. They may also refer to writing assignments. *Prompt* is used in the assessment literature and is a general term meaning specific instructions given to students before they write. *Tasks* may be the same as a prompt, or, in the task-based literature, may refer to something that can be described according to a number of parameters such as reasoning demands, amount of planning time, or whether or not specific content is provided. *Genre* can refer to something as specific as a restaurant review or a business letter, but in the literature discussed here, it refers to more general types of writing, such as narrative, argumentative, or descriptive.

Teaching Tip

- Give students a variety of writing tasks and genres.

Written Language Development in Instructional Contexts

A common area of research in the L2 writing literature is how students' writing develops over the course of some instructional period. Studies examining development may focus on specific linguistic features, or they may use various measures of complexity (syntactic or lexical), accuracy, or fluency (called *CAF* or *CALF* measures). Which measures to use, however, has been a matter of debate for years. One of the first discussions of *CALF* measures was Wolfe-Quintero, Inagaki, and Kim (1998). Most of the studies that they reviewed examined correlations between the *CALF* measures and quality measures such as holistic ratings, or between *CALF* measures and proficiency level, so these studies did not examine development during an instructional period. A few studies, however, were longitudinal and were able to document changes over time on some measures. Wolfe-Quintero et al.'s conclusions were limited because of the lack of reliability reported, the different methods used to place students into levels in the cross-sectional studies, and the paucity of longitudinal studies. One interesting finding, however, was that accuracy measures correlated better with holistic measures of essay quality than with external proficiency measures, such as in-house tests used to place students into levels. This finding is related to the lack of evidence that accuracy changes much in instructional contexts (Polio & Shea, 2014; Yoon & Polio, 2017), at least in one-semester studies.

Key Concept

CALF measures: This stands for complexity, accuracy, lexical, and fluency measures, sometimes called CAF, with lexical measures considered as part of complexity. Some of these constructs, such as accuracy, are easier to define, but complexity may be seen as having different dimensions, such as sentence length, the number of dependent clauses, or the length of noun phrases. The most appropriate measure for all of the constructs has been debated in the literature at some point. Similar measures are used in oral language research but some are specific to written language.

Teaching Tip

- Don't be discouraged if students keep producing errors.

A collection of studies in Connor-Linton and Polio (2014) examined students' language development over the course of English for Academic Purposes (EAP) writing classes. The studies used various measures to try to document changes in constructs such as syntactic complexity (Bulté & Housen, 2014), accuracy (Polio & Shea, 2014), and clusters of discourse features (Friginal & Weigle, 2014). Bulté and Housen found change in some of the complexity measures, while Polio and Shea found almost none in the accuracy measures. Friginal and Weigle were the most successful of the three studies in documenting change. They used Biber's multidimensional approach (Biber, 1988, 1995, 2006), which studies clusters of features related to different genres of discourse. For example in their study, essays exhibited features related to a personal focus (e.g., second person pronouns, *that*-deletion, causative verbs) at the beginning of the semester and then changed to a more informational focus (e.g., agentless passives, prepositions, and concrete nouns).

Another key study is Verspoor, Schmid, and Xu (2012), a cross-sectional study that examined 64 measures across five levels of proficiency of Dutch school-age learners of English. As expected, they found different measures of complexity, accuracy, and fluency increased differently across different levels. Taken together, these studies are useful in choosing measures to use in experimental studies, but none has given us a yardstick for measuring progress in L2 writing classes or programs. Whether or not a common measure is attainable is a matter of debate; oral language measures have not been fully successful at universal measures of language learning, with perhaps the exception of the work by Pienemann (2007; Pienemann & Kefßler, 2012), who was able to propose stages based on general non-language-specific features. With the amount of variation in literacy skills of language learners and individual differences in the resources learners draw on as they write, including explicit knowledge and the time they take to focus on language and correct or revise, such a yardstick may not be possible.

Empirical Evidence

In this section we detail various empirical studies that mirror each of the four issues in the last section. For each topic, there are many more studies that could be included, but we have chosen to foreground well-designed, recent studies.

Charlene Polio and Jongbong Lee

The Role of Writing in SLA

There is surprisingly little research that directly examines how writing can facilitate SLA. Earlier, we mentioned Laufer and Hulstijn's (2001) involvement load hypothesis, which suggests that if students use new vocabulary in a written task, they will retain the vocabulary better. Keating (2008) conducted a study of Spanish learners to determine if writing would help students retain vocabulary better than in a reading or cloze activity. He found that the group that wrote sentences using the new words best retained them. However, when he factored in the time that it took to complete the tasks, the writing group did not perform any better. Huang, Willson, and Eslami (2012) conducted a 12-study meta-analysis of the involvement load hypothesis by examining studies that compared an output with a no-output condition. They found that output led to retention and that composition writing provided the greatest gains when only one output task was used. They too could not rule out time-on-task as a factor. Thus, although writing helps learners retain vocabulary, we do not know if it is the composing process itself or the additional time spent with the vocabulary. The studies related to the involvement load hypothesis focused on vocabulary, and we have even less direct information about how writing might facilitate the learning of morphosyntactic structures.

One of the few studies that directly addressed the modality of instruction was Kim's (2008) small-scale study of five- and six-year-old ESL students. Kim varied the types of instruction that the students received between only oral input and production, and integrated oral and written production in a type of time-series study. For example, in the integrated instruction, students wrote journals about a story instead of discussing the story. She showed that students performed better on oral assessments after the integrated instruction than after the oral-only instruction. The assessments were based on semantic, pragmatic, and grammatical acceptability as well as on the amount of oral language production.

Corrective Feedback

A variety of studies that have claimed to show the effectiveness of corrective feedback have been criticized for not controlling group variables such as amount of instruction (Bitchener, Young, & Cameron, 2005) or frequency of student writing (Chandler, 2003). Two more recent studies showing the positive effects of feedback, however, were well designed. Hartshorn et al. (2010) studied 28 students in a treatment group and 19 in a contrast group in a 15-week intensive university ESL class. Students in the treatment group wrote almost every day for 10 minutes. They received feedback with coded symbols and had to rewrite until all the errors were gone. The students in the contrast group wrote four multidraft papers and received different types of feedback, including feedback on errors. The study was designed to test a feedback method called dynamic corrective feedback, based on skill acquisition theory (see DeKeyser, 2007). Hartshorn et al. found that the treatment group made significantly fewer errors while their rhetorical competence, fluency, and complexity scores did not suffer. This study is noteworthy because it was done with intact classes showing that such an intensive treatment could be given in a classroom setting.

Using a different design, Van Beuningen, De Jong, and Kuiken (2012) conducted a tightly controlled experimental study in which 134 Dutch secondary students, 80% of whom were L2 learners, were randomly assigned to one of four groups: direct feedback (errors corrected); indirect feedback (errors coded); self-correction (time given to self-correct); and additional writing (time spent on a new writing task). Students wrote and were given comprehensive

feedback on one piece of writing. They then produced a new text one week later and again four weeks later. Van Beuningen et al. found that both corrective feedback groups wrote more accurately than the two groups who did not receive feedback and that they did not use simpler language so as to avoid errors. Direct feedback was most effective for grammatical errors and indirect feedback for nongrammatical (lexical and spelling) errors.

These two studies differed greatly in the length of the treatment (15 weeks vs. one time) and in terms of student populations, yet they both showed a positive effect for feedback. Evans, Hartshorn, and Strong-Krause (2011) replicated Hartshorn et al. (2010) with matriculated university students and found the same results. Both Bitchener and Knoch (2015) and Polio and Park (2016) have called for the Van Beuningen et al. (2012) study to be replicated with a longitudinal design to determine if the effects are durable.

There has been enough research on written corrective feedback to warrant at least three meta-analyses focusing only on written feedback. In the first meta-analysis, Truscott (2007) excluded single-treatment designs, such as the design used in Van Beuningen et al. (2012), and he concluded that there was a small negative effect for error correction. Kao and Wible (2014) used Truscott's (2007) inclusion criteria but were able to include 26 studies because of the additional research conducted after 2007. They found positive effects for feedback as did Kang and Han (2015), who included 22 studies in their meta-analysis.

Shintani and Aubrey (2016) expanded the scope of feedback research by examining synchronous computer-mediated feedback. Japanese EFL students wrote texts that elicited the hypothetical conditional. In addition to a standard feedback and no-feedback group, one group received immediate feedback as they wrote in a Google Docs environment, and this group improved the most on their use of the hypothetical conditional on a new piece of writing, most likely because of the immediateness of the feedback. Taken together, studies on corrective feedback suggest that in some circumstances, there is a positive effect for corrective feedback.

Teaching Tip

- Correct errors on some assignments and have the students revise those assignments.

The Effect of Different Tasks on Written Language Production

Many researchers have manipulated both task conditions and complexity features of writing prompts to examine language production. Planning time, for example, was manipulated in Ellis and Yuan (2004) and Ong and Zhang (2010). Ellis and Yuan measured fluency in terms of syllables per minute and the number of dysfluencies (i.e., crossed out and changed words). Ong and Zhang measured fluency in terms of total number of words and words per minute. Ellis and Yuan found that planning significantly affected fluency and syntactic variety but not complexity or accuracy. Ong and Zhang also found increased fluency in the planning condition as well as increased lexical complexity, but they did not test for syntactic complexity or accuracy. It appears then that planning can affect fluency but the effects on other features of writing are less clear.

Kuiken and Vedder (2008) had students write on two prompts about choosing a vacation spot. The prompts differed in terms of the number of elements that the students had to consider (e.g., location, breakfast, swimming facilities). They found that the complex

condition elicited more accurate writing but not more complex writing. Other studies that targeted syntactic complexity also found no effect for task complexity on linguistic complexity (Kormos, 2011, 2014). Tavakoli (2014) studied task complexity operationalized by the number of storylines in a set of pictures that students had to describe: the more complex task had two storylines that needed to converge. Tavakoli found no significant differences in the syntactic complexity measures between the two written tasks, but she did find differences when the story narratives were oral. These studies show that although planning time may affect some aspects of written language, the effects of task complexity are not as robust in written language as they are in oral language.

In contrast, studies on genres have revealed robust differences in the complexity of learner writing. Both Lu (2011) and Yoon and Polio (2017) found that students' language was more complex in argumentative writing than in narratives. Interestingly, one study of task differences that found language difference across tasks was Frear and Bitchener (2015), but they operationalized differences, in part, based on the presence of reasoning demands, one of the differences between argumentative (+ reasoning demands) and narrative essays (– reasoning demands). Frear and Bitchener found that tasks with the reasoning demands contained more adverbial clauses but not adjectival or noun clauses. Both Yoon and Polio and Frear and Bitchener suggested that their results might be due to differences in the necessary communicative functions for the task and not the cognitive load in performing the task. Yoon and Polio drew on Biber and Conrad (2009) to show how the differences in complexity are related to Biber and Conrad's characterization of argumentative and narrative texts, while Frear and Bitchener drew on Ryshina-Pankova and Byrnes (2013), who noted features of academic register that could account for task differences.

Written SLA in Instructional Contexts

One type of longitudinal written SLA studies are those conducted in a study abroad context (for a review, see Sasaki, 2011). One recent study, Godfrey, Treacy, and Tarone (2014), compared learning French abroad and at home. The researchers used not only the American Council on the Teaching of Foreign Languages (ACTFL) scale but also measures of complexity, accuracy, and fluency to assess the students' writing. They studied eight university students of French, four of whom studied abroad and four of whom studied at home. We single out this study because they described what was happening in the French classes in the US in comparison to study abroad. Because of the small sample size, it is difficult to draw clear conclusions, but it appeared that both groups made some improvements but on different measures. In this type of study, it is difficult to know what was due to instruction and what was due to exposure in the target culture. While this may seem obvious, we point out that studies of ESL classes suffer from the same problem (e.g., those in Connor-Linton & Polio, 2014).

We highlight here four studies that examined written language in a variety of instructional contexts including secondary EFL classes in a Dutch high school (Verspoor & Smiskova, 2012), an intensive British EAP program (Mazgutova & Kormos, 2015), a Japanese EFL class (Yasuda, 2011), and a fourth-grade science class in the US (De Oliveira & Lan, 2014). In addition, they each represent a different approach to examining written language development and instruction, yet they all, to some extent, try to link development to what is happening in the classroom.

Verspoor and Smiskova (2012) followed 20 Dutch high school students studying English for two years. Half of the students were in a low-input group studying English for two hours a

week, while half were in a high input group studying 15 hours a week. The study focused on the use of chunks or formulaic sequences with the hypothesis that the high-input group would use more chunks in their writing. Verspoor and Smiskova included structures such as lexical collocations (*apologize profusely*), compounds (*forest fire*), textual prepositions (*with respect to*), and textual adverbs (*in other words*), among others. For some types of chunks, but not many, the high-input group produced more in their writing. By then focusing on one student from each group, Verspoor and Smiskova were able to document different trajectories with regard to chunk use; the high-input group student's variability decreased over the course of the study, that is, the ratio of chunks per 100 words became more stable.

Key Concept

Chunks or sequences: These refer to pieces of language that may be learned or processed together. They include idioms (*to rain buckets*), but also figurative language (*I'm starving to death*), phrasal verbs (*to pick up*), lexical collocations (*to muster courage*), discourse markers (*first of all*), or words that simply tend to co-occur (*let me know*).

Teaching Tip

- Get students to focus on genre-specific chunks or formulas.

Mazgutova and Kormos (2015) examined language development in a 4-week intensive EAP writing class. Students were given feedback but no explicit language instruction. The authors examined change on a variety of syntactic complexity (such as mean length of T-unit) and lexical diversity measures. In addition, they coded the essays for features that have been claimed to be features of academic writing such as conditional clauses and complex noun phrases, both of which increased in the lower proficiency group. Both the low and high groups improved their lexical diversity despite no explicit vocabulary instruction. The most striking finding is that both groups used a smaller variety of syntactic structures at the end of the course. Mazgutova and Kormos suggested that development might not be linear or that students begin limiting their writing to structures that are more prominent in academic writing at some level of proficiency. The students in their study were placed at the B2 and C1 levels on the Common European Framework of Reference scale, which categorized them as “independent” or “proficient” users of the language, respectively.

De Oliveira and Lan (2014) and Yasuda (2011) both examined writing development for students taught using a genre-based approach. Yasuda (2011) taught a university EFL class in Japan that revolved around writing different types of emails. Among the various studies of development in instructional contexts, this one describes in the most detail what was happening in the class in terms of the syllabus design, procedures, and feedback. Although there was no control group, Yasuda compared pre- and post-course email tasks. In terms of language development, she found that the students were able to write longer texts but that their lexical diversity (i.e., the use of a greater variety of words) did not change over the 13 weeks. She found, however, that the students increased their lexical sophistication, usually measured by considering the frequency of words (i.e., the less frequent the more sophisticated) but in this

Charlene Polio and Jongbong Lee

study measured by the frequency of genre-specific expressions (e.g., *I want to* vs. *I would be grateful if*). In this way, the measure of sophistication linked development to instruction.

De Oliveira and Lan (2014) took a case study approach to studying a fourth grade ESL student learning to write science texts. They had the student write a text about a science experiment before the teacher implemented a specific type of genre-based instruction (based on Martin & Rose, 2003) that involves deconstructing a target text and then jointly constructing a text with the student before the student then constructs a text alone. De Oliveira and Lan were able to trace features of the teacher's focus on language to the student's text, specifically the use of field-specific vocabulary and the use of a wider variety of temporal connectors. On one hand, the study focused on only one student, but on the other hand it is the clearest study in terms of relating instruction to written texts.

Teaching Tip

- Scaffold with students as they construct new genres.

Pedagogical Implications

Writing Activities in General Language Courses

The evidence that students focus on language as they write, both alone and together, is overwhelming, even if the long-term effects of this focus have not been well documented. This focus suggests that writing-to-learn-language activities should be used in most language classes. For example, teachers can have students write to describe a picture that later might be used for an oral information gap activity. Alternatively, students can write after an oral role play, perhaps by narrating a description of the event from the role play. The former activity will give students a chance to search for necessary vocabulary while the latter will allow them to reuse and hopefully internalize language from the oral activity. One commonly used writing activity in the literature is a dictogloss (e.g., Kowal & Swain, 1994; Swain & Lapkin, 2001). In this activity, students listen to a passage and then try to reconstruct it, usually writing together. Prince (2013) discusses the related research and variations on the dictogloss task.

In general, teachers can be creative about writing activities and do not need to limit themselves to what might be viewed as real-life writing tasks. Another example of a pedagogic writing task that encourages students to focus on language is a story continuation task described by Wang and Wang (2014). They had students read a story in either English or Chinese and then continue the story in English. They found that students produced fewer errors after reading the English version and used some of the vocabulary from the English version.

Varying Tasks and Genre

While it is not completely clear that more complex tasks result in more complex language, it is clear that students use more complex language in certain genres and that they need to learn genre-appropriate language. Thus, genres and tasks should be varied even for beginners. If students keep to writing assignments that elicit simple language, they may not have an opportunity to develop their language. However, as shown in De Oliveira and Lan (2014), students need appropriate scaffolding as they write new genres. There are various

ways to scaffold, but the model from Martin and Rose (2003) is one way. In this model, the teacher helps students deconstruct a model text with regard to language and organizational features. The teacher then helps students write a text before they have to do it alone. Another option is to have students construct texts together as much of the research on collaboration has shown that students can help each other produce and revise texts (for a comprehensive discussion, see Storch, 2013).

Taking a Middle Ground on Corrective Feedback

There is now enough evidence that some types of corrective feedback can be helpful. Of course, there are also studies suggesting that it is not always effective. Thus, a middle ground needs to be taken. Because simply producing written output can facilitate acquisition, teachers should not avoid having students write because they do not have time to give feedback. For example, on a dictogloss, teachers can have students compare their texts to the original as the teacher walks around the room and answers questions; there is no need to collect the texts and correct every error. Some teachers have voiced opposition to not correcting errors, but Ferris (2014) surveyed 129 writing teachers who taught a range of students including both native speakers and multilingual writers and found that only 37% commented on all or most of students' writing. This finding suggests that because the majority of teachers have students do some writing for which they are not given feedback, opinions about the necessity of corrective feedback may be changing.

Future Directions

As with many areas of ISLA, research on writing in languages other than English is sparse, particularly at the advanced levels. Thus, replicating any of the studies discussed here in new contexts with different languages might prove interesting, particularly with languages that use non-Roman scripts. In addition, like other areas of ISLA, more longitudinal research needs to be conducted. Specifically, we know something about the effects of different prompts on learner language, but we do not know about the effects of implementing various writing tasks over the course of an instructional period. In addition to these two overarching suggestions, we include specific directions related to technology, interventions not related to error correction, and research on direct links between instruction and SLA.

Technology

We did not discuss the role of technology in instructed written SLA, but refer to some studies to highlight this as an area for further study. Kessler, Bikowski, and Boggs (2012) investigated web and project-based collaborative writing activities, and analyzed group collaborative writings in Google Docs. They found that L2 learners had a tendency to engage more with meaning than form and to make more correct grammatical changes than ungrammatical ones. Simultaneous editing and scaffolding offered by partners in such activities may help L2 learners. However, it is not clear if web-based collaborative writing activities are more beneficial than offline collaborative writing.

Technology affords innovative types of feedback as discussed by Shintani and Aubrey (2016). Taking a different approach to the use of technology, Elola and Oskoz (2016) investigated teacher feedback given via Word versus oral feedback with screencast software. They found that the students preferred oral feedback for global issues and written feedback for

Charlene Polio and Jongbong Lee

form problems. In the end, however, their revisions did not differ significantly. Two areas that might be the focus of future research are related to automated feedback (e.g., Lavolette, Polio, & Kahng, 2015) and the effects of corpus consultation, a topic not addressed here, but see Benavides (2015) and Liu and Jiang (2009) for the effects on students' written language.

Interventions That Push Complexity Development

We cited studies that found students used more complex language in certain genres. What would be useful to know is how consistent use of more complex language affects long-term development; we know of no studies that attempt to push development by having students produce more complex language over an extended period of time. For example, while there are studies of language that students use in dictoglosses, we do not know of any extended intervention studies that use them to teach specific structures. For example, one could construct a series of dictoglosses that focus on various morphosyntactic structures. Students completing the dictoglosses could be compared to a control group who write using their own language, as opposed to reconstructing texts, for increased linguistic complexity.

Linking Instruction to SLA

Here we return to De Oliveira and Lan (2014) and Mazgutova and Kormos (2015) as examples of research that we would like to see extended. Mazgutova and Kormos's study was interesting because they documented progress after only 4 weeks of an intensive EAP class. We do not know exactly what caused the progress, so a similar study but supplementing the data with observations and student and teacher interviews would be helpful. In contrast, De Oliveira and Lan (2014) examined what happened in the classroom but presented data from only one student. Their study might benefit from additional quantitative data by examining the writing from a larger number of students. Put another way, a mixed methods study drawing on both quantitative and qualitative data from a course or a set of lessons might better help us link instruction and writing development.

Conclusion

There is no doubt that writing and some types of interventions related to writing instruction can help learners focus on language even if the long-term effects are not obvious. We hope that this chapter has provided convincing evidence showing that writing instruction does not have to be limited to contexts such as academic writing classes and can benefit all learners. In contrast, for students who do need to learn specific genres related to their learning goals, teacher feedback, scaffolding, and collaborative activities can facilitate language learning.

References

- Belcher, D., & Hirvela, A. (2008). Introduction. In D. Belcher & A. Hirvela (Eds.), *The oral-literate connection* (pp. 1–8). Ann Arbor, MI: University of Michigan Press.
- Benavides, C. (2015). Using a corpus in a 300-level Spanish grammar course. *Foreign Language Annals*, 48, 218–235.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

- Biber, D. (1995). *Dimensions of register variation: Across-linguistic perspective*. Cambridge: Cambridge University Press.
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5, 97–116.
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press.
- Bigelow, M., Delmas, R., Hansen, K., & Tarone, E. (2006). Literacy and the processing of oral recasts in SLA. *TESOL Quarterly*, 40, 665–689.
- Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. New York/London: Routledge.
- Bitchener, J., & Knoch, U. (2015). Written corrective feedback studies: Approximate replication of Bitchener & Knoch (2010a) and Van Beuningen, De Jong & Kuiken (2012). *Language Teaching*, 48, 405–418.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191–205.
- Brooks, L., & Swain, M. (2009). Linguaging in collaborative writing: Creation of and response to expertise. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 58–89). New York: Routledge.
- Bruton, A. (2009). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing*, 18, 136–140.
- Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, 12, 267–296.
- Connor-Linton, J., & Polio, C. (2014). Comparing perspectives on L2 writing: Multiple analyses of a common corpus: Introduction. *Journal of Second Language Writing*, 23, 1–9.
- Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, 7, 482–511.
- DeKeyser, R. (2007). *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press.
- DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94–112). New York: Routledge/Taylor & Francis.
- De Oliveira, L. C., & Lan, S.-W. (2014). Writing science in an upper elementary classroom: A genre-based approach to teaching English language learners. *Journal of Second Language Writing*, 25, 23–39.
- Dulay, H. C., & Burt, M. K. (1973). Should we teach children syntax? *Language Learning*, 23, 245–258.
- Dulay, H. C., & Burt, M. K. (1974). Natural sequences in child second language acquisition. *Language Learning*, 23, 245–258.
- Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26, 59–84.
- Elola, I., & Oskoz, A. (2016). Supporting second language writing using multimodal feedback. *Foreign Language Annals*, 49, 58–74.
- Evans, N. W., Hartshorn, K. J., & Strong-Krause, D. (2011). The efficacy of dynamic written corrective feedback for university-matriculated ESL learners. *System*, 39, 229–239.
- Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form vs. content. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 178–190). Cambridge: Cambridge University Press.
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1–10.
- Ferris, D. (2014). Responding to student writing: Teachers' philosophies and practices. *Assessing Writing*, 9, 6–23.

Charlene Polio and Jongbong Lee

- Frear, M., & Bitchener, J. (2015). The effects of cognitive complexity on writing complexity. *Journal of Second Language Writing*, 30, 45–57.
- Friginal, E., & Weigle, S. (2014). Exploring multiple profiles of L2 writing using multi-dimensional analysis. *Journal of Second Language Writing*, 26, 80–95.
- Godfrey, L., Treacy, C., & Tarone, E. (2014). Change in French second language writing in study abroad and domestic contexts. *Foreign Language Annals*, 47, 48–65.
- Guénette, D. (2007). Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16, 40–53.
- Gutiérrez, X. (2008). What does metalinguistic activity in learners' interaction during a collaborative L2 writing task look like? *Modern Language Journal*, 92, 519–537.
- Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11, 329–350.
- Hartshorn, K., Evans, N., Merrill, P., Sudweeks, R., Strong-Krause, D., & Anderson, N. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, 44, 84–109.
- He, L., & Shi, L. (2012). Topical knowledge and ESL writing. *Language Testing*, 29, 443–464.
- Housen, A., & Pierrard, M. (2005). Instructed second language acquisition: Introduction. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 1–26). Berlin: Mouton de Gruyter.
- Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta-analytic study. *Modern Language Journal*, 96(4), 544–557.
- Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, 63, 330–367.
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, 99, 1–18.
- Kao, C. W., & Wible, D. (2014). A meta-analysis on the effectiveness of grammar correction in second language writing. *English Language Teaching & Learning*, 38, 29–69.
- Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, 12, 365–386.
- Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. *Language Learning & Technology*, 16, 91–109.
- Kim, Y. (2008). The effects of integrated language-based instruction in elementary ESL learning. *Modern Language Journal*, 92, 431–451.
- Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, 20, 148–161.
- Kormos, J. (2014). Differences across modalities of performance. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 193–216). Amsterdam: John Benjamins.
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3, 73–93.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, 17, 48–60.
- Lalande, J. F. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, 66, 140–149.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1–26.
- Lavolette, E., Polio, C., & Kahng, J. (2015). The accuracy of computer-assisted feedback and students' responses to it. *Language, Learning & Technology*, 19, 50–68.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85–102.
- Lillis, T., & Curry, M. J. (2010). *Academic writing in a global context: The politics and practices of publishing in English*. New York: Routledge.

- Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93, 61–78.
- Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing*, 30, 66–81.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York/London: Routledge/Taylor Francis.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie & T.K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45, 36–62.
- Manchón, R. (2011). Situating the learning-to-write and writing-to-learn dimensions of L2 writing. In R. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 3–14), Amsterdam: John Benjamins.
- Martin, J.R., & Rose, D. (2003). *Working with discourse: Meaning beyond the clause*. London: Continuum.
- Matsuda, P.K. (2015). Identity in written discourse. *Annual Review of Applied Linguistics*, 35, 140–159.
- Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for academic purposes programme. *Journal of Second Language Writing*, 29, 3–15.
- Ong, J., & Zhang, L. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing*, 19, 218–233.
- Ortega, L. (2011). Reflections on the learning-to-write and writing-to-learn dimensions of second language writing. In R. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 237–250). Amsterdam: John Benjamins.
- Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition*, 6, 186–214.
- Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, 10, 52–79.
- Pienemann, M. (2007). Processability theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 137–154). Mahwah, NJ: Lawrence Erlbaum.
- Pienemann, M., & Kessler, J. (2012). Processability theory. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 228–247). New York: Routledge.
- Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, 47, 101–143.
- Polio, C. (2012a). Replication in published applied linguistics research: An historical perspective. In G. Porte (Ed.), *Replication in applied linguistics: A practical guide* (pp. 47–91). Cambridge: Cambridge University Press.
- Polio, C. (2012b). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21, 375–389.
- Polio, C., & Park, J. (2016). Language development in second language writing. In R.M. Manchón & P. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 287–306). New York: Routledge.
- Polio, C., & Shea, M.C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10–27.
- Prince, P. (2013). Listening, remembering, writing: Exploring the dictogloss task. *Language Teaching Research*, 17, 486–500.
- Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20, 83–95.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–58.
- Ryshina-Pankova, M., & Byrnes, H. (2013). Writing as learning to know: Tracing knowledge construction in L2 German compositions. *Journal of Second Language Writing*, 22, 179–197.

- Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45, 81–105.
- Semke, H. D. (1984). Effects of the red pen. *Foreign Language Annals*, 17, 195–202.
- Shi, L. (1998). Effects of prewriting discussions on adult ESL students' compositions. *Journal of Second Language Writing*, 7, 319–345.
- Shintani, N., & Aubrey, S. (2016). The effectiveness of synchronous and asynchronous written corrective feedback on grammatical accuracy in a computer-mediated environment. *The Modern Language Journal*, 100, 296–319.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Storch, N. (2011). Collaborative writing in L2 contexts: Processes, outcomes, and future directions. *Annual Review of Applied Linguistics*, 31, 275–288.
- Storch, N. (2013). *Collaborative writing in L2 classrooms*. Bristol: Multilingual Matters.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step toward second language learning. *Applied Linguistics*, 16, 371–391.
- Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogical tasks* (pp. 99–118). London: Longman.
- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research*, 37, 285–304.
- Tardy, C. (2012). A rhetorical genre theory perspective on L2 writing development. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 165–190). Boston, MA/Berlin: De Gruyter Mouton.
- Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning: Insights from and for L2 writing* (pp. 217–236). Amsterdam: John Benjamins.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327–369.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16, 255–272.
- Van Beuningen, C., De Jong, N., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62, 1–41.
- Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21, 239–263.
- Verspoor, M., & Smiskova, H. (2012). Foreign language writing development from a dynamic usage based perspective. In R. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 17–46). Boston, MA: De Gruyter Mouton.
- Wang, C., & Wang, M. (2014). Effect of alignment on L2 written production. *Applied Linguistics*, 36, 503–526.
- Weigle, S. (2014). Considerations for teaching second language writing. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (pp. 222–237). Boston, MA: National Geographic Learning.
- Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: Writing vs. speech. *Learning and Instruction*, 10, 37–53.
- Weissberg, B. (2006). *Connecting speaking and writing*. Ann Arbor, MI: University of Michigan Press.
- Williams, J. (2005). *Teaching writing in second and foreign language classrooms*. New York: McGraw-Hill.
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21, 321–331.

- Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Honolulu: University of Hawai'i Press.
- Xu, C. (2009). Overgeneralization from a narrow focus: A response to Ellis et al. (2008) and Bitchener (2008). *Journal of Second Language Writing*, 18, 270–275.
- Yasuda, S. (2011). Genre-based tasks in foreign language writing: Developing writers' genre awareness, linguistic knowledge, and writing competence. *Journal of Second Language Writing*, 20, 111–133.
- Yoon, H., & Polio, C. (in press). Linguistic development in two genres. *TESOL Quarterly*.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Section IV

Instructed Second Language Acquisition Learning Environments



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

ISLA in East Asian Contexts

Yuko Goto Butler

Background

Theory building and empirical investigations of second language acquisition (SLA) have been largely led by researchers in selected English-speaking countries such as the US and UK. In the last couple of decades, however, a growing number of researchers have started conducting studies in other regions, including East Asia. In subscribing to the premise of SLA that “the context of learning does not alter the cognitive mechanisms that drive learning” (Loewen, 2015, p. 144), researchers from different regions have contributed to our understanding of cognitive mechanisms in second language (L2) learning. But others have questioned some of the assumptions of universal L2 development and the widely accepted pedagogical approaches for facilitating L2 acquisition; instead, they have highlighted the importance of the role of context in understanding SLA (e.g., Littlewood, 2007; Prabhu, 1990).

In this chapter, focusing on instructed SLA (ISLA) research in East Asia, I consider how contexts can influence our conceptualization of some of the key notions in SLA as well as the pedagogical approaches and strategies for facilitating SLA. Major contributions of ISLA in East Asia can be summarized as (1) challenging universal and cognitive-centered approaches to conceptualizing SLA; (2) searching for contextually appropriate and effective pedagogical approaches for language teaching; and (3) productively researching language learning amid changing learning spaces and learner characteristics.

The first contribution is to challenge the universal assumptions regarding what to develop. ISLA research in East Asia has questioned how some of the critical concepts in SLA are defined, and has proposed incorporating social dimensions in conceptualizing SLA or broadening the definitions to allow some flexibility in interpretation. For example, major concepts such as *communicative competence*, *learner autonomy*, and *motivation* were once mainly conceptualized as cognitive states that reside in individuals. However, socially oriented views of these notions have gained increasing recognition (e.g., social dimensions in language testing in McNamara & Rover, 2007; social dimensions in learner autonomy in Murray, 2014; the contextualized model of motivation in Dörnyei, 2003). Studies on ISLA in East Asia often either serve as a driving force for such reconceptualization or provide empirical support to justify the conceptual modifications.

The second contribution of ISLA research in East Asia is its search for pedagogical approaches and strategies for developing target language proficiency that are most effective in a given context. For example, as I will discuss in detail, certain pedagogical approaches such as *communicative language teaching* (CLT) and *task-based language teaching* (TBLT) have been promoted worldwide, including in East Asia, but many reports indicate that these approaches, in their original forms, do not necessarily work well across contexts. When new pedagogical approaches and strategies were introduced in East Asia, local educators gradually adapted them to suit their own contexts and students (e.g., Butler, 2011; Thomas & Reinders, 2015). As will be discussed, in the process of the adaptation, local teachers and researchers reconceptualized or shed new light on the traditional approaches and incorporated them in the newly introduced pedagogies, instead of totally discarding the existing practices. As a result, both local educators and researchers often develop localized hybrid models of language teaching.

Finally, ISLA studies in East Asia actively deal with issues resulting from rapid and constant changes of learning spaces and learners. For example, substantial learning is now taking place outside of traditional classrooms in digital learning spaces. Learner characteristics are increasingly diversified as well. Young learners and lifelong learners are growing in number. These changes in learning spaces and learner characteristics challenge static and categorized conceptualizations in SLA (e.g., second language vs. foreign language; native vs. nonnative; language education vs. content language integrated learning [CLIL]; first language vs. second language; communicative language vs. academic language; virtual vs. real; explicit teaching vs. implicit teaching). Nonstatic, dynamic, and fluid models and approaches in SLA are increasingly called for.

While a growing number of studies in East Asia have examined ISLA on a variety of languages, the research on ISLA in this region has so far predominately been concerned with English language learning/teaching. Thus, I draw examples mostly from English learning/teaching. Despite my focus on English, many of my arguments in this chapter apply to any other language learning/teaching. Geographically, the chapter primarily concerns China (including Hong Kong), Japan, South Korea, and Taiwan because most studies are from these areas, but it also includes new research coming out of Thailand and Vietnam.

The organization of this chapter is as follows. First, I provide some background on ISLA research in East Asia and summarize early major contributions. I then discuss each of the major theses just outlined. I conclude with suggestions for future research and pedagogical implications.

Socioeducational Contexts in East Asia

Loewen (2015) argued that ISLA concerns situations in which learners are attempting to acquire a target language in the midst of “some systematic attempt to manipulate the conditions for learning” (p. 5). The systematic manipulation can be achieved either by manipulating the linguistic input (e.g., using authentic texts or modified texts according to the learners’ proficiency levels) or by manipulating the process in which the learners engage with the input (e.g., asking students to pay attention to a certain linguistic forms or to read the text for pleasure). But one can also argue that both means of systematic manipulation are influenced by the socioeducational contexts where learning and teaching are taking place. I add one more dimension in which socioeducational context has an influence: one’s attitudes about learning and teaching.

First, how the linguistic input is manipulated, both quantitatively and qualitatively, varies depending on the socioeducational context. When it comes to English teaching/learning, East Asia has been commonly considered a foreign-language context as opposed to a second-language context. Traditionally, *second-language (L2) learning* refers to learning a language other than one's first language (L1) that takes place in a community where the target language is dominantly spoken (e.g., English learning by immigrants in the US). *Foreign language (FL) learning* refers to cases in which non-L1 education is carried out where the target language is not the primary language spoken (e.g., English learning in Japan). As a result, L2 learners are more likely exposed to a greater amount of the target language input compared with FL learners. Not only the amount of input but also the types of input that learners can receive often differ between L2 and FL contexts. In formal classroom settings, FL learners may receive the majority of their input from teachers who speak a variety (or varieties) of the target language. Their input may be largely modified for instructional purposes and sequentially presented in accordance with local, predefined curriculum. The range of language use in classrooms may be limited to certain domains in FL contexts (e.g., Longcope, 2009; Sato, 2010).

It is undeniable that English learning in East Asia is largely conducted in somewhat limited English environments both quantitatively and qualitatively, and that educators in East Asia have made tremendous efforts to maximize the amount and the type of input in English in their classrooms. The promotion of teaching English using only English (to the exclusion of the L1)—an approach that is spreading widely in East Asia (Dearden, 2014)—can be considered one such effort.

It is important to note, however, that the distinction between L2 learning and FL learning is increasingly becoming fuzzy as greater opportunities to receive input outside of the formal classroom setting are available, at least for some learners. Thanks to advances in technology, a growing number of learners in East Asia have greater opportunities to be exposed to English of various types without leaving their home countries. As a result, we see widening gaps in access to the input in the target language, not only across communities but also within a community, according to learners' socioeconomic status and digital literacy skills. A growing number of children have started learning a FL, English in particular, at an earlier age, and their exposure to the target language is diversified across their age, socioeconomic status, and contexts; some young learners may learn the target language in a bilingual immersion program or a *content language integrated learning (CLIL)* program where the learners learn select academic subjects through the target language (Butler, 2015a). *Shadow education* (e.g., private tutoring and learning taking place at cram schools, English conversation schools) is a massive industry in East Asia, and some learners receive a tremendous amount of instruction outside of their formal schooling (Bray & Lykins, 2012). Study-abroad programs are gaining popularity, starting from the primary school levels (e.g., Song, 2011 for a case in South Korea), and learners can move back and forth between a traditionally defined "L2 context" and "FL context." The *native speaker fallacy*, or *native-speakerism* (Holliday, 2005)—a belief that a certain native speakers' input should be the model—remains powerful in many parts of East Asia (Chan & Evans, 2011). However, language educators are gradually beginning to question such goal setting (e.g., Chan, 2013; Miyagi, Sato, & Crump, 2009), and what can be called *postnative models* have been explored (e.g., English as a lingua franca in Asia: see Kirkpatrick, 2006). Therefore, in order to reflect such changing educational environments and growing criticisms of the L2-versus-FL dichotomous conceptualization, I characterize English-learning/teaching in East Asia as "L2/FL" (implying a fuzzy boundary between the two) for the rest of the discussion in this chapter.

Second, the ways that learners engage in and make use of language input are influenced by their sociocultural traditions, educational environments, and learning needs. In East Asia, English is often far more than a means of communication. It also plays a critical role as a major academic subject in the exam-driven educational systems. High achievement in English is considered a sign of academic success and diligence, and explicit form-focused instruction¹ such as grammar-translation activities have been dominant in East Asia in order for students to perform well on high-stakes exams. Under such socioeducational contexts, when new pedagogical approaches/methods that focus on meaning rather than forms, such as CLT and TBLT, were introduced in the formal school curricula in a top-down manner sometime around the late 1980s (CLT) and late 1990s (TBLT), researchers found tremendous gaps between the policy intention and the actual implementation in classrooms. According to Butler (2011), difficulties in implementing CLT and TBLT reported by earlier studies (studies published up to the mid-2000s) can be attributed to three major factors: (1) conceptual constraints (e.g., mismatch between learning concepts underlying CLT and the traditional East Asian values of learning and teaching;² teachers' lack of understanding of CLT); (2) classroom-level constraints (e.g., lack of human resources and materials, large classroom sizes, limited instructional hours, and classroom management issues); and (3) societal/institutional-level constraints (e.g., grammar-translation oriented high-stakes tests and limited opportunities to use English in daily life). Similar challenges and problems with implementing TBLT have been reported as well.

These earlier responses to the top-down implementations of CLT and TBLT were significant in that they questioned some of the implicit assumptions in SLA: the assumptions that critical notions such as “communicativeness” mean the same thing irrespective of contexts and that pedagogical approaches and methods that have proven to be effective in one context should work in other contexts. At the same time, however, these studies often emphasized “problems” with the implementations while tending to overlook positive aspects of new approaches, and some studies fell into stereotypical discussions of East versus West (e.g., East Asian students prefer passive learning). Such generalizations appear to be based on a static and uniform notion of East Asian education, and can easily mislead or mask the reality that education is highly complicated and diverse within East Asia (Butler, 2011; Littlewood, 1999). Moreover, as Lai (2015) stated, such generalizations “will not help the field move forward” (p. 24).

Third, socioeducational contexts also influence people's perceptions of and attitudes toward learning and teaching. Take motivation for language learning as an example. For decades, Gardner's socioeducational model (Gardner, 1985), which originated in Canada, has been very influential. However, researchers who worked in English as a FL or lingua franca started questioning the role of *integrative motivation*, the model's major construct, in the context of English as a lingua franca (ELF). Integrative motivation refers to learners' desire to integrate into or be part of the target language community. For learners of ELF, the target language communities may not be clearly identifiable in the first place. For example, Yashima (2002) collected data from English learners in Japan and proposed that *international posture*—one's desire to communicate internationally—turned out to be a more significant motivation for the students. Similarly, a number of researchers in East Asia (e.g., Kang, 2005, from Korea; Koga, 2010, from Japan; Peng & Woodrow, 2010, from China) paid attention to learners' willingness to communicate (MacIntyre, Clément, Dörnyei, & Noels, 1998), meaning learners' preparedness to use the target language when opportunities are available. These studies on willingness to communicate all indicated the importance of classroom context factors—such as cooperation among learners, teacher immediacy, and

task-oriented instruction—as greatly influencing learners’ willingness to communicate in their target language.

In sum, in East Asia, L2/FL learning, and English learning in particular, has distinct roles (both as an academic pursuit in the formal education system and as a means for communication). When new concepts and pedagogical approaches/strategies in SLA were imported from the West, researchers found that educators faced difficulties in implementing them in their classrooms. By articulating the problems and conflicts, those earlier studies made us aware of the important role that socioeducational contexts play in language learning. However, stereotyped approaches that characterize East Asian contexts, such as arguments based on Confucianism, are potentially misleading. Socioeducational contexts in East Asia are in the midst of drastic changes and are increasingly diversified (e.g., Chan & Rao, 2009; Hannum, Park, & Butler, 2010). L2/FL teaching needs to acknowledge and respond to such changes.

Current Issues and Empirical Evidence

A growing number of researchers in East Asia indicate that learners in East Asia are responsive to innovative SLA pedagogies, such as TBLT, if the pedagogies are adapted for local contexts. A number of case studies present hybrid pedagogical models that incorporate reconceptualized local traditions. Furthermore, nontraditional learning spaces, most notably in the use of technology, are on the rise. In this section, I examine examples of these new trends in SLA in East Asia.

Task-Based Language Teaching (TBLT)

Following CLT, TBLT rests on the notion that language learning is much more than an acquisition of structural and lexical knowledge. Rather, learners’ communicative competence is developed through meaningful interaction, namely *tasks*. While theorists generally agree on the basic notion of tasks as activities involving the use of language that is focused on meaning (Skehan, 2003), a more precise definition remains somewhat controversial (e.g., Ellis, 2003; Van den Branden, 2006). This lack of clarity in conceptualizing tasks is partly responsible for creating confusion among teachers in East Asia (Butler, 2011, 2015a; Littlewood, 2007). After observing difficulties with implementation in East Asian classrooms, Littlewood (2007, 2014) proposed a five-level model of tasks to guide teachers. The model classifies tasks along a continuum of activities, from form-focused to meaning-focused, as follows: (1) noncommunicative learning, (2) precommunicative language practice, (3) communicative language practice, (4) structured communication, and (5) authentic communication. Littlewood’s model attempted to move beyond a dichotomy of noncommunicative exercises versus communicative tasks to “a loose conceptual framework” (Littlewood, 2014, p. 360), while not assuming that there is a single effective communicative teaching method that every teacher should follow. By doing so, Littlewood suggested that teachers can explore different types of tasks in their communicative classrooms according to their own professional experiences, their students’ needs, and various contextual factors.

Unlike the stereotypical image of East Asian students as passive and unlikely to actively respond or speak up in class, a number of studies have indicated that students were positive about communicative tasks (e.g., Chung & Huang, 2009, from Taiwan; Hood, Elwood, & Falout, 2009 from Japan; Nguyen, Newton, & Crabbe, 2015 from Vietnam). Teachers’ attitudes appeared to influence their practice as well as their students’ engagement in tasks.

For example, Nishino (2012) from Japan found that high school teachers' perception of their students' needs and their ability to work in pairs and groups influenced how teachers implemented tasks. In Iwashita and Li's (2012) study of Chinese college students, instructors' positive attitudes toward communicative tasks led to students' frequent interaction and active participation in tasks despite unfavorable conditions for task implementation, which included large class sizes and students' unfamiliarity with the tasks.

Researchers in East Asia are also increasingly interested in identifying how best to integrate or adapt TBLT in local contexts instead of simply describing difficulties to adopting TBLT. Studies on task adaptation in East Asia have been mostly case studies employing classroom observations and interviews with stakeholders (i.e., teachers and students) and primarily concerned with identifying conditions and elements that make the TBLT implementation workable and effective in the given context. These studies tend to support a weaker version of TBLT (also referred to as *task-supported language teaching*), where learners are allowed to use tasks to analyze the language, rather than a stronger version of TBLT that advocates subconscious learning through tasks and therefore requires the syllabus to be exclusively composed of tasks (Adams & Newton, 2009). Flexibility in implementation seems to be the key. Compared with what is typically suggested by TBLT methodologists, teachers in East Asia often appear to have much greater involvement in their students' task activities through all phases of TBLT, from planning tasks, assisting students during the tasks, and carrying out posttask activities (e.g., Lingley, 2006 for a Japanese university; Darasawang, 2015; Watson Todd, 2006 for Thai universities). Considering that many teachers in Hong Kong have relied heavily on a traditional *presentation-practice-production* (P-P-P) approach, and that P-P-P has "perceived pragmatic advantages," Carless (2009, p. 64) suggested that instead of completely discarding P-P-P, it may be possible to incorporate P-P-P into TBLT—as long as the limitations of P-P-P (e.g., learners may be able to use target forms and expressions during the lesson but may not be able to acquire them in the long run) are minimized.

There are at least two critical issues with implementing TBLT in East Asia. As secondary school teachers in Hong Kong in Carless (2007, 2009) nicely articulated, the two issues are: (1) placing greater emphasis on grammar instruction in TBLT, and (2) situating TBLT in such a way that students' exam requirements are considered. How best to address these issues, however, remains an unsolved challenge.

The first challenge is to figure out how best to incorporate form-focused instruction into TBLT. Under relatively limited input conditions typically found in East Asia, educators believe that explicit form-focused instruction is indispensable. Indeed, a meta-analysis found that explicit instruction in general is more effective than implicit instruction (Norris & Ortega, 2000).³ While TBLT methodologists generally agree that some sort of form-focused elements need to be incorporated in TBLT, they disagree over when and how to carry it out. Some researchers, such as Willis (1996), have suggested that form-focused elements should be introduced, if necessary, at the posttask phase in order to avoid turning authentic communicative activities into predefined grammar and lexical exercises. This recommendation, however, appears to be counterintuitive for some teachers, those who are used to P-P-P in particular. In East Asia, as seen in Lingley's (2006) study in a Japanese university, teachers generally prefer to conduct explicit form-focused activities at the pretask phase in TBLT because of the students' needs (e.g., students may not be able to perform a given task without practicing pre-identified words and forms) and other institutional requirements (e.g., curriculum defines what to acquire in a given lesson).

The extent to which the introduction of form-focused activities at the pretask phase is effective remains unclear, however. Nguyen et al. (2015) examined instructors' practice of pretask phases at a Thai high school where four-staged tasks (pretasks, rehearsal in pairs, task performance as a form of public display, and posttask activities)⁴ were institutionally implemented. They found that the instructors' practice at the pretask phases varied according to their beliefs; half of the instructors introduced controlled practice activities (e.g., providing useful linguistic expressions and modeling) whereas the other half did not. Interestingly, the majority of the students preferred not to have any form-focused activities at the pretask phases because they did not want to be constrained by predefined vocabulary and expressions. As Nguyen et al. (2015) acknowledged, however, this result might have been largely due to the fact that the participating students had high motivation and proficiency. Indeed, a study conducted among beginning-level Japanese students of English indicated that form-focused pretask planning did help students produce the target form more accurately (Mochizuki & Ortega, 2008). In any event, understanding the students' needs seems to be critical when considering the timing and the strategies for incorporating form-focused elements in TBLT. There may be no one-size-fits-all solution.

The second challenge concerns how best to situate TBLT and communicative task-based assessment in a highly exam-oriented educational system where exams, norm-referenced entrance exams in particular, have had tremendous influence over teaching and learning practice (e.g., Carless, 2011; Littlewood, 2007). Because the content of exams and the procedures used to administer them often strongly regulate the way teachers feel they should teach, teachers may "feel powerless" when it comes to making decisions on teaching and assessment (Hamp-Lyons, 2007, p. 498). Efforts have been made in recent years to change the educational system throughout East Asia, such as the growing tendency to incorporate large-scale proficiency tests that include oral assessment (e.g., Japan) or performance-based assessment (e.g., Hong Kong and South Korea) for admission or placement purposes. However, the intended positive *washback effects* (such as having tests or assessments influence individual and societal educational practices) have not necessarily been observed. This is not surprising, however, because washback effects are the results of a complicated interplay among multiple factors, including the degree of support provided to stakeholders and societal attitudes toward exams (Cheng, Watanabe, & Curtis, 2004). Again, teachers' perceptions of and beliefs toward exams are often more influential over their practice than the actual exam-related constraints; the exam reforms may have limited effects unless teachers have sufficient understanding and receive support for the change (Cheng, Sun, & Ma, 2015; Cheng et al., 2004).

While a good deal of research has been conducted on the design and implementation of tasks (e.g., Adams & Newton, 2009; Thomas & Reinders, 2015), task-based assessment (TBA) has been explored less extensively in East Asia. As Long and Crookes (1992) suggested more than two decades ago, in theory, TBA should be conducted "by way of task-based criterion-referenced tests" (p. 45). However, in practice, a number of issues must be clarified: (1) how task-based criteria should be defined (e.g., based on linguistic performance or task completion); (2) how tasks should be selected for assessment in order to correspond to the criteria (e.g., based on constructs or work samples); and (3) how learners' performance should be evaluated reliably and validly (Butler, 2011). In any of these processes, teachers are expected to play substantial roles.

Hong Kong's school-based assessment (SBA) is one of the few innovative approaches implemented in an "exam-oriented educational system" on a large scale. In 2005, SBA was introduced as part of the Hong Kong Certificate of Education Examination (HKCEE), a very

high-stakes exam certifying students' completion of secondary school education that plays a key role in admission to higher education. SBA was a classroom-based TBA, and it was set to account for 15% of students' total scores on the HKCEE. In SBA, teachers need to assess their students' English performance against defined criteria, while adjusting assessment tasks according to individual students' proficiency levels. SBA was composed of oral interactive tasks and oral presentation tasks designed to build upon individual students' self-selected texts/readings prior to the exam (Davison, 2007). Not too surprisingly, a number of concerns were raised initially, including concerns with respect to the teachers' ability to assess their students' performance and the issue of fairness. After receiving sufficient training and supports, however, the teachers gained confidence in doing SBA (Davison & Hamp-Lyons, 2010). Although HKCEE was replaced by a new exam called the Hong Kong Diploma of Secondary Education (HKDSE) in 2012, an SBA component continues to be part of HKDSE, with some adjustments (Hong Kong Examinations and Assessment Authority, n.d.).

Hong Kong's SBA is definitely a promising move, but it also exemplifies a complicated reality. For example, Luk (2010) examined students' interaction during group discussions in SBA and found that the students made a "collective attempt to present a best impression of themselves as well as the whole group through ritualized, institutionalized, and colluded talk" (p. 46). The students did not challenge others while taking mechanical turns in order to make sure that everybody had an equal chance to talk. Such behaviors resulted in inauthentic interaction. Luk's study showed the difficult and complex nature of task-based assessment if it were meant to play two contradictory roles—namely, showcasing one's best performance and realizing authentic communication (Butler, 2011).

Key Concepts

School-based assessment (SBA): An innovative, task-based assessment practice implemented as part of the Hong Kong Certificate of Educational Examination (HKCEE) and the HKCEE's replacement, the Hong Kong Diploma of Secondary Education. Conducted by teachers, SBA aims to "enhance the validity of the public assessment" and to "include a variety of learning outcomes that cannot be assessed easily through public examinations" (Hong Kong Examinations and Assessment Authority, 2013, p. 1).

Learner autonomy: Originally defined as "the ability to take charge of one's own learning" (Holec, 1981, p. 3), learner autonomy is considered an important individual capacity influencing the process and outcome of language learning. In East Asia, more socially oriented conceptualizations of learner autonomy have been proposed. Such modified conceptualizations include Littlewood's (1999) two levels of autonomy: proactive autonomy and reactive autonomy. According to Littlewood, while proactive autonomy is well aligned with Holec's definition of autonomy, reactive autonomy may be less independent and does not necessarily require complete control over one's own learning.

Reconceptualized "Traditional" Methods and Strategies

With the promotion of CLT and TBLT in East Asia, "traditional" pedagogical methods such as the grammar-translation method and the audio-lingual method have been strongly criticized for failing to effectively develop communicative competence (oral skills in

particular), having excessive teacher control and relying heavily on memorization and repetition. Despite such repeated criticisms, these traditional methods are still popular in East Asia, and it seems that they can still turn out highly proficient, successful learners. For example, Ding (2007), in an interview study with winners of a nationwide college English-speaking competition in China, indicated that these successful English learners found text memorization and imitation to be “the most effective methods of learning English” (p. 271) because these methods “enabled them to attend to and learn collocations and sequences, to borrow these sequences for productive use, to improve pronunciation, and to develop the habit of attending to details of language in the context of language input” (p. 271).

Recently, some researchers have started questioning the traditional-versus-communicative dichotomy, and have advocated treating them as complementary while also shedding new light on the traditional methods (Beaumont & Chang, 2011; Griffiths, 2011; Jin & Cortazzi, 2011). For example, taking a new approach to the grammar-translation method, Lee, Schallert, and Kim (2015) examined the effect of translation as a means of reading instruction on Korean middle school students’ grammatical knowledge and compared its effectiveness with another means of reading instruction—namely, extensive reading. Importantly, the translation activity in their study was different from the typical grammar-translation approach in which teachers explain preselected grammar rules first and then ask students to translate a text applying the grammar rules. In Lee et al. (2015), the teacher did not explicitly explain any preidentified grammar rules nor ask the students to apply the targeted rules in their translation activity. However, in the translation condition, after the students worked on translation individually or in pairs, they could ask the teacher for help with their translation difficulties. In the extensive reading condition, the students individually read books of their choice and wrote short response notes in Korean after each reading. The result indicated that both translation and extensive reading approaches showed score gains in grammar tests; however, unlike the extensive reading, from which higher-proficiency students had more benefit, the translation worked best among mid-proficiency students.⁵ In addition, the students in the translation group, irrespective of their proficiency levels, all showed higher perceived improvement of their general linguistic skills and more positive attitudes toward the activity (e.g., more enjoyment and engagement) than the students in the extensive reading condition. The positive results from their translation activity may be in part due to the fact that it allowed students to have more autonomy and opportunities to interact both with the teacher and peers than the traditional grammar-translation approach.

Innovative approaches to the grammar-translation method have been tried in writing instruction as well. For example, in action research in an English composition class at a Korean college, Kim (2011) used translation to facilitate students’ reflection and collaboration in class. Current SLA pedagogy has strongly promoted process-oriented writing, in which writing processes are emphasized rather than the accuracy of the end product. However Kim saw limited application of the process-oriented pedagogies among her low-proficiency students (e.g., no improvement was made after repeated revisions, the students showed little awareness of their own writing for improvement, etc.). Kim then decided to take advantage of the students’ translation skills, and asked the students to translate what they wrote in English into Korean (by either translating their own or their peers’ English compositions) as a way to become aware of their own problems and to facilitate collaboration in class. Interestingly, the act of translation made the students aware of the importance of

grammar (accuracy of writing) in addition to the importance of the process of writing. Kim's study suggested that translation can allow students, at least those with lower proficiency, to use their L1 as a resource when learning an L2.

Despite being criticized as an unfortunate legacy of audio-lingualism, *imitation* and *repetition* also have survived in language classrooms in East Asia. The role of repetition in child first language acquisition is well recognized. Clark (2003) indicated that linguistic repetition has at least two important functions for children: it "connote[s] acceptance or ratification of the adult terms" and "offer[s] children an opportunity to try to produce the target term in a recognizable fashion and thus practice the as-yet unfamiliar term" (p. 321). Researchers in language socialization also have paid substantial attention to the role of repetition in the course of children's linguistic and sociocultural development and have examined its various practices in socialization, such as revoicing, prompting, guided repetition, and language play across different communities (Moore, 2011).

Recently, researchers reconceptualized the role of imitation and repetition in L2/FL development as well. From an information-processing point of view, repetition is understood as "a resource which not only offers access to new language forms, but also enables learners to proceed from controlled language use to more spontaneous and automatic production" (Pirainen-Marsh & Alanen, 2012, p. 2826). Complexity theory sees "the innovative role of repetition" if it is considered not as merely copying but as "iteration that generates variation" (Larsen-Freeman, 2012, p. 207). In complexity theory, repeating words and utterances always creates new meanings, and this iteration serves as a starting point of the next iteration. This process is considered to be the very act of learning, in which learners adapt their resources to fit a new context. Practitioners may recognize such an innovative role based on their practical experience, as we can see in a statement made by a Chinese language teacher in Marton, Dall'Alba, and Tse (1996): "In the process of repetition, it is not a simple repetition. Because each time I repeat, I would have some new idea of understanding, that is to say I can understand better" (p. 81). Sociocultural theory also sees imitation as a key contribution for development, including language development. Imitation occurs when a child engages in a task that is beyond what he or she is capable of doing independently. In other words, imitating another individual indicates the child's reachable ability level if he/she receives assistance from others. Imitation serves as a bridge for internalizing the intellectual activity through interacting with capable others (Vygotsky, 1978). Finally, Cook (2000), from a sociolinguistic point of view, also stated that

our examination of play suggests that activities often associated with a focus on form (such as repetition, rote learning, and structural analysis and manipulation) *can* take on personal and social significance, and both draw attention to the language, and be "interesting and relevant."

p. 172, emphasis in original

As pedagogical strategies of imitation and repetition, *reciting*, *reading-aloud*, *repeated-reading*, and *shadowing* appear to be popular in language classes in East Asia. Dahlin and Watkins (2000) found that both Hong Kong and German secondary school students indicated that their parents did not impose recitation on them during their childhood, but that Hong Kong school teachers more strongly promoted recitation than their German counterparts. The stronger emphasis on repetition in early school education in Hong Kong may be in part related to Hong Kong's L1 literacy education, which requires students to

master a large number of Chinese characters (Kember, 1996). Reading aloud and repeated reading are often introduced, with the aim of reinforcing graphemic-phonemic correspondences and developing oral fluency (Gibson, 2008). Automatizing phonological coding is particularly challenging for learners whose L1 has a different orthographic system from that of the target language, especially in an input-poor context, such as those experienced by many English learners in East Asia. Once learners can automatize the phonological processing, they can devote more cognitive resources to comprehension. Indeed, Gorsuch and Taguchi (2008) reported that repeated reading helped their Vietnamese college students improve not only their fluency but also comprehension in their English reading. Similarly, shadowing (tracking incoming speech, repeating it, and monitoring the verbalization) is considered beneficial for forming phonological representation of the target language. By eventually automatizing phonological processes through shadowing, learners can efficiently access conceptual representations. In other words, shadowing is not simply a mechanical repetition but involves high cognitive and metacognitive processes and is considered a useful pedagogic strategy in input-poor contexts (Kadota, 2012). In recent years shadowing has gained substantial attention, particularly in Japan, and a growing number of studies—both at the behavioral and neurophysiological levels—have indicated its effectiveness in learners’ listening comprehension, prosody and phonemic development, oral fluency, and the acquisition of formulaic expressions (see Kadota, 2012 for a review of such studies). According to Murphey (2001), shadowing can be characterized according to three continua: (1) from silent to out loud; (2) from complete to selective; and (3) from noninteractive to interactive. In examining students’ shadowing of their conversational partners’ utterances, Murphey observed various types of conversational adjustments and negotiations in their interactive shadowing, which in turn can lead to language development. Considering that shadowing is “a global macro-discursive strategy for language acquisition that many learners can use in many ways” (Murphey, 2001, p. 143), the optimal way to use it depends on the learning/teaching context, the partner, and the purpose of learning.

While repetition/iteration can be used as a pedagogical strategy, it is perhaps most effective when it is implemented in combination with more communicative-oriented methods, as with other “traditional” pedagogical approaches. Moreover, repetition may need to be initiated by learners to be effective. Butler (2015b) asked young learners (ages 11–12) of English in Japan to identify the most effective vocabulary learning strategies and to design, in groups, computer-based English vocabulary learning games while incorporating the self-identified strategies in their game design. She found that the children were aware that repetition was an important element for their English vocabulary learning; however, they also wanted to have control over their repetition activities. The children’s peer evaluation of their game designs showed that they highly valued instructional games that allowed learners to decide what to repeat, when to repeat, and how to repeat. This is quite different from the traditional teacher-initiated repetition activities such as “repeat after me.”

The Expansion of Autonomous Learning Spaces Beyond Classrooms

There are growing opportunities across East Asia (and elsewhere) for learning languages beyond the traditional classroom space (e.g., Benson & Reinders, 2011; Richards, 2015). There is no doubt that the internet and other types of technology significantly enhance

nontraditional learning spaces. If a learner wishes, he or she can access a massive amount of learning resources through computers and mobile devices.

Heift and Chapelle (2012) addressed three core issues that advances in computer technology have brought to the field of SLA. First, they point out that we need to expand our understanding of “interaction” from face-to-face interaction (which SLA research has largely assumed) to something broader that includes computer–human interactions and human–human interactions in virtual spaces. Indeed, we make and negotiate meanings in those new types of interactions, but the ways we do so are often different from traditional face-to-face interactions. In East Asia, interplays between sociopragmatic factors (i.e., face-keeping, politeness, and other pragmatic strategies) and task and technology environments during interaction are topics of great interest among researchers (e.g., Peterson, 2006; Zheng, Young, Wagner, & Brewer, 2009; also see Park, 2008, for a theoretical discussion of this topic). Second, computer technology makes us realize how important it is to pay greater attention to individual differences. We gradually discover that there are substantial individual variations in how learners learn through technology, which in turn can influence their learning rates and developmental trajectories (e.g., Heift, 2008). And third, the increase of computer-mediated language learning beyond the classroom directly speaks to the issue of learner autonomy (e.g., Benson & Reinders, 2011); SLA researchers need to better understand how learners develop autonomous strategies to improve their learning. While technology affords the most notable out-of-school learning spaces, we also see growing opportunities in East Asia to use the target language through various physical autonomous learning spaces such as self-access centers (e.g., *English Corners* in China in Gao, 2008; *English Café* in Japan in Murray, Fujishima, & Uzuka, 2014) and study-abroad programs.

Learner autonomy has gained much attention among researchers in East Asia in recent years (e.g., Griffiths et al., 2014; Murray, 2014). Learning through computer technology and other learning beyond the classroom can be institutionalized or led by a teacher, but such approaches are largely led by learners themselves. While a cognitive-based definition of learner autonomy—“the ability to take charge of one’s own learning” (Holec, 1981, p. 3)—has been long accepted, researchers in East Asia observe “various manifestations of autonomy” in social learning spaces (Murray, 2014, p. 242) and advocate a reconceptualization of autonomy. To capture such variations of autonomy based on his long-term observations in East Asia, Littlewood (1999) proposed *reactive autonomy*, distinguished from *proactive autonomy*. Proactive autonomy aligns well with Holec’s (1981) definition of autonomy, affirming learners’ individuality and self-directed control over contexts. Different from proactive autonomy, reactive autonomy is “the kind of autonomy which does not create its own directions but, once a direction has been initiated, enables learners to organize their resources autonomously in order to reach their goal” (Littlewood, 1999, p. 75). One can characterize reactive autonomy as a more socially oriented notion of autonomy because it can be achieved in social settings through interdependency and collaboration (e.g., an initial direction or task may be set by a teacher). Importantly, reactive autonomy can be considered to be a precondition for developing proactive autonomy, but it also can be treated as a goal in its own right.

Broadened conceptualizations of learner autonomy such as Littlewood’s were welcomed by researchers in East Asia, who often have difficulties characterizing their students’ “autonomous” behaviors in ways that align with the original definition. For example, Yasuhima’s (2014) Japanese high school students’ involvement in a model United Nations project and Murray et al.’s (2014) *English Café* participants at a Japanese university were not

autonomous in the original sense (as defined by Holec, 1981) but can be characterized by these new variations.

Future Directions

ISLA research in East Asia has a number of promising directions for exploration. The first is its potential contribution to theory building, by incorporating social dimensions into SLA theories. ISLA research in East Asia has highlighted the role that context plays in language learning and teaching. As we have seen, for example, researchers in East Asia have reported a number of challenges with respect to implementing CLT and TBLT. There surely are local specificities in their challenges; however, as Lai (2015) suggested, certain contextual elements may be more commonly observable when innovation is employed. Accumulating more information and systematically analyzing the process of contextualizing innovation in wider contexts can provide useful information for theorizing social dimensions in SLA.

Related to the preceding point, we perhaps need to revisit the widely accepted assumption in SLA that “context of learning does not alter the cognitive mechanisms that drive learning” (Loewen, 2015, p. 144) and to systematically examine if this assumption really holds true. So far, we have limited research in East Asia that directly investigates the role of context over one’s cognition or cognitive processing related to L2/FL acquisition. However, we have seen that contexts influence the way in which the input that learners receive is manipulated both quantitatively and qualitatively, the manner in which learners engage with varying input, and learners’ affective and emotional variables in language learning. Moreover, such contextual influences are constantly changing and evolving. Considering these findings, it appears to be worth investigating if or how contexts interact with one’s cognitive mechanisms.

Lastly, we need to better capture changing environments and individual differences. As with any other scientific inquiry, SLA research as a field often categorizes various behaviors and phenomena. While such categorizations are helpful for grasping a general picture of the behaviors and phenomena, they may overly simplify the reality. As we discussed, boundaries used in SLA research, such as L2-versus-FL and communicative-versus-noncommunicative, are increasingly fuzzy. Similarly, grouping learners into certain types or making arguments based on the average behaviors (e.g., mean test scores) may mask the dynamics of individual differences. There is no question that the field is in need of methodological innovations and creative approaches to more fully capture a dynamic reality.

Pedagogical Implications

ISLA studies in East Asia have brought clear practical implications; it is critically important to adopt flexible pedagogical approaches while taking the contextual factors and learners’ characteristics and needs into account. Ecologically valid pedagogical approaches are necessary. More specific suggestions can be found in the following Teaching Tips. ISLA studies in East Asia also often make us aware of hidden universal assumptions behind theories. Scheffler (2012) nicely summarized the critical point: “An important goal of SLA theory is to explain how learning is accomplished through teaching. No teaching procedure claimed by teachers to be effective should be disregarded, even if the focus of current theory is on something completely different” (p. 604). In the end, practitioners’ wisdom drives the theory.

Teaching Tips

- What works best is largely determined in context. Whatever pedagogies are used, be flexible in implementing them while paying careful attention to contextual factors and learner needs. Sometimes, local wisdom can be effectively incorporated into new methods.
- Traditional teacher-led, form-focused methods such as the grammar-translation method seem to have limited effect. However, such methods may support students' L2/FL learning when they are implemented in such a way that they (1) embrace students' autonomy; (2) facilitate collaboration and interaction among students as well as between teachers and students; (3) are combined with more communicative oriented methods; and (4) meet the needs of the students.
- Be responsive to changing environments and learners' needs. Make use of various opportunities to learn the target language—both in class and outside of the class—to help learners develop their autonomy.

Notes

1. The term “form-focused instruction” has been used inconsistently in previous studies. In this chapter, I adopt Ellis and Shintani's (2014) definition: “instruction that involves some attempt to focus learners' attention on specific properties of the L2 so that they will learn them” (p. 337); form-focused instruction can include different types of instruction (e.g., explicit instruction and implicit instruction).
2. Examples of such values include the traditional Confucian norms, such as learning being the acquisition of knowledge that primarily resides in books and teachers being a possessor and messenger of such knowledge. A common discourse in East Asia claimed that such values would not cope well with oral-focused and student-centered instruction, which were believed to be major premises in CLT in East Asia.
3. Norris and Ortega's (2000) meta-analysis was conducted predominantly among studies on adult learners. We know little about the case among young learners.
4. It is interesting to see the inclusion of rehearsal and public display as part of the tasks in Nguyen et al. (2015). This task format can be considered a local modification.
5. The proficiency levels were determined by both the students' grades in English in the previous semester when the study was conducted as well as by two additional in-house general English proficiency tests covering all four skills.

References

- Adams, R., & Newton, J. (2009). TBLT in Asia: Constraints and opportunities. *Asian Journal of English Language Teaching*, 9, 1–17.
- Beaumont, M., & Chang, K.-S. (2011). Challenging the traditional/communicative dichotomy. *ELT Journal*, 65(3), 291–299.
- Benson, P., & Reinders, H. (Eds.). (2011). *Beyond the language classroom*. New York: Palgrave Macmillan.
- Bray, M., & Lykins, C. (2012). *Shadow education: Private supplementary tutoring and its implications for policy makers in Asia*. Mandaluyong, Philippines: Asian Development Bank.
- Butler, Y.G. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics*, 31, 36–57.
- Butler, Y.G. (2015a). English language education among young learners in East Asia: A review of current research (2004–2014). *Language Teaching*, 48(3), 303–342.

- Butler, Y. G. (2015b). The use of computer games as foreign language learning tasks for digital natives. *System*, 54, 91–102.
- Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System*, 35, 595–608.
- Carless, D. (2009). Revisiting the TBLT versus P-P-P. *Asian Journal of English Language Teaching*, 19, 49–66.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-Heritage settings*. New York: Routledge.
- Chan, C.K.K., & Rao, N. (Eds.). (2009). *Revisiting the Chinese learner: Changing contexts, changing education*. Hong Kong/Dordrecht, NL: The Comparative Education Research Centre, University of Hong Kong/Springer.
- Chan, J.Y.H. (2013). Towards a lingua franca pedagogical model in the Hong Kong classroom: A sociolinguistic enquiry. *Asian EFL Journal*, 15(2), 183–216.
- Chan, J.Y.H., & Evans, S. (2011). Choosing an appropriate pronunciation model for the ELT classroom: A Hong Kong perspective. *Journal of Asia TEFL*, 8(4), 1–24.
- Cheng, L., Sun, Y., & Ma, J. (2015). Review of washback research literature within Kane's argument-based validation framework. *Language Teaching*, 48(4), 436–470.
- Cheng, L., Watanabe, Y., & Curtis, A. (Eds.). (2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- Chung, I.-F., & Huang, Y.-C. (2009). The implementation of communicative language teaching: An investigation of students' viewpoints. *Asia-Pacific Education Research*, 18(1), 67–78.
- Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.
- Cook, G. (2000). *Language play, language learning*. Oxford: Oxford University Press.
- Dahlin, B., & Watkins, D. (2000). The role of repetition in the processes of memorizing and understanding: A comparison of the views of German and Chinese secondary school students in Hong Kong. *British Journal of Educational Psychology*, 70, 65–84.
- Darasawang, P. (2015). Material design for TBLT in Thailand: Balancing process and content. In M. Thomas & H. Reinders (Eds.), *Contemporary task-based language teaching in Asia* (pp. 279–290). London: Bloomsbury.
- Davison, C. (2007). Views from the chalkface: English language school-based assessment in Hong Kong. *Language Assessment Quarterly*, 4(1), 37–68.
- Davison, C., & Hamp-Lyons, L. (2010). The Hong Kong certificate of education: School-based assessment reform in Hong Kong English language education. In L. Y. Cheng & A. Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 248–266). New York: Routledge.
- Dearden, J. (2014). *English as a medium of instruction: A growing global phenomenon*. London: British Council.
- Ding, Y. (2007). Text memorization and imitation: The practices of successful Chinese learners of English. *System*, 35, 271–280.
- Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning*, 53(1), 3–32.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R., & Shintani, N. (Eds.). (2014). *Exploring language pedagogy through second language acquisition research*. London: Routledge.
- Gao, X. (2008). The “English corner” as out-of-class learning activity. *ELT Journal*, 63, 60–67.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitude and motivation*. London: Edward Arnold.
- Gibson, S. (2008). Reading aloud: A useful learning tool? *ELT Journal*, 62(1), 29–36.
- Gorsuch, G., & Taguchi, E. (2008). Repeated reading for developing reading fluency and reading comprehension: The case of EFL learners in Vietnam. *System*, 36, 253–278.
- Griffiths, C. (2011). The traditional/communicative dichotomy. *ELT Journal*, 65(3), 300–308.
- Griffiths, C., Oxford, R. L., Kawai, Y., Kawai, C., Park, Y. Y., Ma, X., Meng, Y., & Yang, N. (2014). Focus on context: Narratives from East Asia. *System*, 43, 50–63.

- Hamp-Lyons, L. (2007). The impact of testing practices on teaching: Ideologies and alternatives. In J. Cummins & C. Davison (Eds.), *The international handbook of English language teaching* (Vol. 1, pp. 487–504). Norwell, MA: Springer.
- Hannum, E., Park, H., & Butler, Y. G. (Eds.). (2010). *Globalization, demographic change and educational inequality in East Asia* (Research in Sociology in Education, Vol. 17). London: Emerald.
- Heift, T. (2008). Modeling learner variability in CALL. *Computer-Assisted Language Learning*, 21(4), 305–321.
- Heift, T., & Chapelle, C. A. (2012). Language learning through technology. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 555–569). London: Routledge.
- Holec, H. (1981). *Autonomy in foreign language learning*. Oxford: Pergamon.
- Holliday, A. R. (2005). *The struggle to teach English as an international language*. Oxford: Oxford University Press.
- Hong Kong Examinations and Assessment Authority. (n.d.). School-based assessment. Retrieved from <http://www.hkeaa.edu.hk/en/sba/>
- Hong Kong Examinations and Assessment Authority. (2013). Hong Kong diploma of secondary education examination: Information on school-based assessment. Retrieved from <http://www.hkeaa.edu.hk/en/hkdse/introduction/>
- Hood, M., Elwood, J., & Falout, J. (2009). Student attitudes toward task-based language teaching at Japanese universities. *Asian Journal of English Language Teaching*, 19, 19–47.
- Iwashita, N., & Li, H. L. (2012). Patterns of corrective feedback in a task-based adult EFL classroom setting in China. In A. Shehadeh & C. A. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 137–161). Amsterdam: John Benjamins.
- Jin, L., & Cortazzi, M. (2011). Re-evaluating traditional approaches to second language teaching and learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 558–575). New York: Routledge.
- Kadota, S. (2012). *Shadowing ondoku-to engoshutoku-no kagaku* [Shadowing, oral-reading and English acquisition science]. Tokyo: Cosmopier.
- Kang, S.-J. (2005). Dynamic emergence of situational willingness to communicate in a second language. *System*, 33, 277–292.
- Kember, D. (1996). The intention to both memorize and understand: Another approach to learning? *Higher Education*, 31, 341–354.
- Kim, E.-Y. (2011). Using translation exercises in the communicative EFL writing classroom. *ELT Journal*, 65(2), 154–160.
- Kirkpatrick, A. (2006). Asian Englishes: Implications for English language. *Teaching Asian Englishes*, 9(2), 4–19.
- Koga, T. (2010). Dynamicity of motivation, anxiety and cooperativeness in a semester course. *System*, 38, 172–184.
- Lai, C. (2015). Task-based language teaching in the Asian context: Where are we now and where are we going? In M. Thomas & H. Reinders (Eds.), *Contemporary task-based language teaching in Asia* (pp. 12–29). London: Bloomsbury.
- Larsen-Freeman, D. (2012). On the roles of repetition in language teaching and learning. *Applied Linguistics Review*, 3, 195–210.
- Lee, J., Schallert, D. L., & Kim, E. (2015). Effects of extensive reading and translation activities on grammar knowledge and attitudes for EFL adolescents. *System*, 52, 38–50.
- Lingley, D. (2006). A task-based approach to teaching a content-based Canadian studies course in an EFL context. *Asian EFL Journal*, 8, 122–139.
- Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics*, 20, 71–94.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*, 40, 243–249.

- Littlewood, W. (2014). Communication-oriented language teaching: Where are we now? Where do we go from here? *Language Teaching*, 47(3), 349–362.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26, 27–56.
- Longcope, P. (2009). Differences between the EFL and the ESL language learning contexts. *Studies in Language and Culture*, 30(2), 303–320.
- Luk, J. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly*, 7, 25–53.
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a second language: A situational model of second language confidence and affiliation. *Modern Language Journal*, 82(4), 545–562.
- Marton, F., Dall’Alba, G., & Tse, K. T. (1996). Memorizing and understanding: The keys to the paradox? In D. Watkins & J. B. Biggs (Eds.), *The Chinese learner: Cultural, psychological, and contextual influence* (pp. 69–83). Hong Kong: Comparative Education Research Centre & Melbourne: Australian Council for Educational Research.
- McNamara, T., & Rover, C. (2007). *Language testing: The social dimension*. Oxford: Wiley-Blackwell.
- Miyagi, K., Sato, M., & Crump, A. (2009). To challenge the unchallenged: Potential of non-“standard” Englishes for Japanese EFL learners. *JALT Journal*, 31(2), 261–273.
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12(1), 11–37.
- Moore, L. (2011). Language socialization and repetition. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The Handbook of language socialization* (pp. 209–226). Malden, MA: Wiley-Blackwell.
- Murphey, T. (2001). Exploring conversational shadowing. *Language Teaching Research*, 5(2), 128–155.
- Murray, G. (Ed.). (2014). *Social dimensions of autonomy in language learning*. London: Palgrave Macmillan.
- Murray, G., Fujishima, N., & Uzuka, M. (2014). The semiotics of place: Autonomy and space. In G. Murray (Ed.), *Social dimensions of autonomy in language learning* (pp. 81–99). London: Palgrave Macmillan.
- Nguyen, B. T. T., Newton, J., & Crabbe, D. (2015). Preparing for tasks in Vietnamese EFL high school classrooms: Teaching in action. In M. Thomas & H. Reinders (Eds.), *Contemporary task-based language teaching in Asia* (pp. 170–188). London: Bloomsbury.
- Nishino, T. (2012). Modeling teacher beliefs and practices in context: A multimethods approach. *Modern Language Journal*, 96, 380–399.
- Norris, J., & Ortega, L. (2000). Effectiveness of instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Park, J. (2008). Linguistic politeness and face-work in computer-mediated communication, Part 1: A theoretical framework. *Journal of the American Society for Information Science and Technology*, 59(13), 2051–2059.
- Peng, J.-E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning*, 60, 834–876.
- Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning*, 19, 79–103.
- Piirainen-Marsh, A., & Alanen, R. (2012). Repetition and imitation: Opportunities for learning. In N. M. Seel (Ed.), *Encyclopedia of the science of learning* (pp. 2825–2828). New York: Springer.
- Prabhu, N. S. (1990). There is no best method—why? *TESOL Quarterly*, 24, 161–176.
- Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *REL C Journal*, 46(1), 5–22.
- Sato, R. (2010). Reconsidering the effectiveness and suitability of PPP and TBLT in the Japanese EFL classroom. *JALT Journal*, 32(2), 189–200.

Yuko Goto Butler

- Scheffler, P. (2012). Theories pass: Learners and teachers remain. *Applied Linguistics*, 33(5), 603–607.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36, 1–14.
- Song, J. (2011). Globalization, children's study abroad, and transnationalism as an emerging context for language learning: A new look for language teacher education. *TESOL Quarterly*, 45(4), 749–758.
- Thomas, M., & Reinders, H. (Eds.). (2015). *Contemporary task-based language teaching in Asia*. London: Bloomsbury.
- Van den Branden, K. (2006). Introduction: Task-based language teaching in a nutshell. In K. Van den Branden (Ed.), *Task-based language education: From theory to practice* (pp. 1–16). Cambridge: Cambridge University Press.
- Vygotsky, L. S. (1978). *Mind and society*. Cambridge, MA: Harvard University Press.
- Watson Todd, R. (2006). Continuing change after the innovation. *System*, 34, 1–14.
- Willis, J. (1996). *A framework for task-based learning*. London: Collins.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *Modern Language Journal*, 86(1), 54–66.
- Yashima, T. (2014). Self-regulation and autonomous dependency amongst Japanese learners of English. In G. Murray (Ed.), *Social dimensions of autonomy in language learning* (pp. 60–77). London: Palgrave Macmillan.
- Zheng, D., Young, M. F., Wagner, M. M., & Brewer, R. A. (2009). Negotiation for action: English language learning in game-based virtual worlds. *Modern Language Journal*, 93(4), 489–511.

Study Abroad and ISLA

Carmen Pérez-Vidal

Background

The interest in studying the effects of Study Abroad (SA) as a context for potential personal, linguistic, cultural and academic development seems undeniable and undiminishing—SA has also been referred to as ‘Stay Abroad,’ ‘Residence Abroad’ (Coleman, 2002) or, with a more general meaning, as ‘mobility’ (Jackson, 2010). SA is not a new phenomenon. Indeed, Erasmus of Rotterdam in the 16th century was a mobile scholar. What is new is the upward trend in mobility figures in contemporary society in general, only paralleled by the interest in exploring SA effects within second language acquisition (SLA) research, the main focus of this chapter.

Why such a new trend in mobility these days? The answer is internationalization, which has spread as a new goal, hand in hand with the globalization of the economy (Falk & Kanach, 2000; Jackson, 2013), underlying mobility across the globe, and, clearly so in education (Banks & Bhandari, 2012; DeWit & Merckx, 2012). Paige, Cohen, and Shiveley (2004) attest from the perspective of US programmes: “Study abroad is clearly a global educational phenomenon, a ‘growth industry’ in higher education, and contributes to broader internationalization efforts in colleges and universities” (p. 253). Indeed, according to the Open Doors Report, published by the Institute of International Education, in 2015 there were 4.5 million mobile college and university students worldwide, for which the US remained the destination of choice, with almost double the number hosted by the UK, the second leading host country. Within the US, over 304,467 US students embarked on a SA programme before graduating from college or university during the academic year 2013–2014, representing an increase of 5% over that year.

Furthermore, current figures stretch the impact of SA to the border of employability issues (Leask, 2015; Pérez-Vidal, 2015a; Trooboff & Rayman, 2008). Within Europe, a recent report attests to the higher employability rates of SA students, who are

half as likely to experience long-term unemployment compared to those who did not go abroad [. . .] while, five to ten years after graduation, 70% of previously mobile alumni hold a managerial position compared to only 40% of those who did not go abroad.

European Commission, 2016¹

Carmen Pérez-Vidal

In this case, SA and languages have been part of a wider, long-standing general strategic policy (see for a summary Coleman, 2006; Pérez-Vidal, 2015a) geared toward languages and multilingualism, to promote the existing linguistic diversity in Europe (European Commission, 1995). The European Action Scheme for the Mobility of University Students (ERASMUS) was thus launched in 1987, and since then it has established itself with more than three million students having experienced an exchange abroad (Coleman, 1998; European Commission, 2013). Together with the momentum given by internationalization as a goal, the role of English as an international language, and its increasing popularity as the ‘lingua franca’ medium of instruction, certainly in Europe, has also contributed to SA becoming the popular enterprise that it is today (see Coleman, 2006; Smit & Dafouz, 2013). Research on the use of English as a lingua franca among mobile university students shows that in Europe, 3.4% of courses are English-Medium-Instruction (EMI) (Wächter & Maiwörm, 2014). This is often seen as a way to attract incoming SA international students, and to allow (outgoing) local students to get of a flavour of internationalization from home (Leask, 2015).

Key Concept

Study Abroad and English-Medium-Instruction (EMI): Study Abroad programmes may count on the availability of courses specifically taught for an international student audience. These days English would be the language mostly used on those courses, which have come to be called English-Medium-Instruction (EMI) courses.

Teaching Tip

In this time of growing internationalization in education, programmes may have to resort to using English as a lingua franca for content courses offered to both local and visiting students. EMI may prove useful to overcome the lack of proficiency in the local language(s), which, nonetheless, may also be of interest to visiting students and deserve specific attention. EMI should also generate an international context ‘at home’ for the potential benefit of local student populations.

Turning to research on the impact of SA periods spent in a target language (TL) country, two decades ago a seminal volume was edited (Freed, 1995a) that marked the beginning of what has been identified as the first period of SA research within the field of SLA research (Collentine, 2009; Pérez-Vidal, 2014b). The volume included both European and American studies on linguistic and sociolinguistic effects of SA, measuring gains mostly with broad measurement instruments. Freed’s (1995b) introduction underlined the fact that, at the time, relatively few empirical studies existed that addressed, in a carefully controlled and in-depth manner, the specific question of the linguistic impact of SA, and even fewer contrasted it with formal classroom instruction at home. As Collentine (2009) later emphasized, Freed’s volume was the first attempt at a state of the art account in a subfield of

enquiry dating back to the 1960s. Its major contribution was that it succeeded in “framing SA research within the SLA theory-building enterprise [. . .] as a means of studying the effects of learning context on acquisition” (pp. 219–220), particularly as regards the fact that ‘context matters,’ as the benefits of SA on TL development did not appear to be the same as those of formal instruction (FI) (Regan, 1995).

The term ‘Study Abroad’ was used in those studies to refer to the educational experience of leaving home on a temporary basis to embark on academic programmes abroad that “As a rule, [. . .] combine language and/or content learning in a formal classroom setting, along with immersion in the native speech community” (Freed, 1995b, p. 25), to provide informal (out-of-class) learning. A SA *programme* itself thus requires an organizational apparatus that is provided by the educational institutions, with an administrative, an academic, and a social component.

Throughout the following decades, new themes, besides linguistic impact, and new angles to approach them, have emerged, reflecting the social turn in the field of SLA (Block, 2003), in what has been referred to as the second period of SA research (Collentine, 2009; Pérez-Vidal, 2014b). Such new themes include, following Collentine’s (2009) tripartite distinction: (1) cognitive, psycholinguistic approaches looking into cognitive processing mechanisms displayed while abroad; (2) sociolinguistic approaches analysing input and interaction from a macro- and a micro-perspective; and, most centrally, (3) sociocultural approaches derived from a paradigm shift from a language-centric (i.e., *etic*) approach to a learner-centric (i.e., *emic*) one (Devlin, 2014). Indeed, within such a paradigm, and in order to focus on the learner and his/her immediate circumstances, SA research has recently begun to investigate nonlinguistic individual differences that affect learning in such a context, that is: (1) intercultural sensitivity and identity changes; (2) affects, such as foreign language anxiety (FLA) or willingness to communicate (WTC) and enjoyment; and (3) social networks, particularly through the use of new technologies and social platforms, and their effect on linguistic practice. These topics all have resulted in noteworthy collections of studies and publications (such as Collentine & Freed, 2004a; DuFon & Churchill, 2006; Gore, 2005; Pellegrino, 2005; Pérez-Vidal, 2014a; Regan, Howard, & Lemée, 2009; Tracy-Ventura, Dewaele, Koylu, & McMannus, 2016).

Against such a backdrop, Collentine (2009, p. 219) has clearly identified the challenge for SA research, as a quest to seek to understand the interaction “between [such] cognitive, sociolinguistic and sociocultural factors in the construction of a comprehensive theory of SLA.” Research has gone a long way indeed. A decade ago, according to Collentine and Freed (2004b, p. 164), there was “no evidence that one context of learning is uniformly superior to another for all students, at all levels of language learning, and for all language skills.” Now, as DeKeyser (2014) emphasizes, “a picture is beginning to emerge of what language development typically takes place [during SA] and what the main factors are that determine the large amount of variation found from one study to another” (p. 313).

From an SLA research perspective, there is no doubt these days that language acquisition differs according to the context in which learners find themselves, be it SA or FI, the latter understood as the conventional second/foreign language classroom (see for example Collentine, 2009; Llanes, 2011; Pérez-Vidal, 2014b; Sanz, 2014). Opportunities for practicing the language in the form of amount of exposure and interaction are undoubtedly larger during SA than in FI. This is clearly revealed when analysing their nature in detail. Pérez-Vidal (2014b), based on Kasper and Rose (2002), presents them in sharp contrast, at the two opposite ends of a continuum, by portraying SA on one end, offering “a naturalistic

learning context in which learners are immersed in the TL and culture with potentially massive amounts of sociolinguistically varied input, output and interaction opportunities available to them,” and FI at the other end of the continuum, “with no opportunities to practice the target language outside the classroom” (p. 23), where, as Collentine (2009) further describes: “[Formal instruction] Learning contexts manage input and output so that learners will attend to form and take intentional steps toward improving their linguistic expertise” (p. 218).

However, this situation must be further qualified: the newly explored themes in SA research referred to earlier are already beginning to uncover the differences individual learners show, for example, in their abilities to put into play and make the most out of previous FI, when sojourning abroad, and in their ability to avail themselves of the opportunities a SA context offers them (see Dörnyei & Ushioda, 2009). This new research focus has been fundamental; as Rees and Klapper (2008) note, SA learners must be placed in the foreground and seen first and foremost as individuals. I would argue, how could it be otherwise? SA is an individual, and an often challenging, endeavour. It is ultimately for each individual learner to display the adequate strategies needed to establish contact with TL speakers while abroad (Collentine & Freed, 2004b), in order to practice the language, to benefit from the linguistic landscape, that is the language used on public and commercial signs (Backhaus, 2007), and the local culture (including the media, the arts, sociopolitical events, etc.).

And yet, somewhat paradoxically, recent studies have highlighted that SA learners often struggle to find opportunities for interaction with their native-speaking counterparts (Devlin, 2014; Jackson, 2008; Pellegrino, 2005). The reality seems to be, as Mitchell, Tracy-Ventura, and McManus (2015) have observed, that “the construction of social groups/communities of practice turns out to be easier for many sojourners when getting together with other international students, than with the locals” (p. 8). Now, is this what SA is all about, with sojourners socializing within ‘international circles’ and not local ones?

The answer to this question may well be partly yes, and partly no. As suggested at the beginning of this section, SA need not be an end in itself, but one of the paths toward gaining what Dörnyei and Ushioda (2015) refer to as an international stance, that is, a view of the world that takes into account countries and languages other than the learners’ own one(s), and that often uses English as a lingua franca as the means of communication. As such, SA is however, sometimes seen with sceptical eyes, as being tinted with neoliberal colours, by research that takes a more social and critical standpoint (Block, 2003; Murphy-Lejeune, 2002) concerning the fact that these days, internationalization underlies mobility across the globe, and, as already discussed, clearly so in education (DeWit & Merckx, 2012).

Key Concept

Study Abroad Objectives: A period spent in a TL country can fulfil different objectives. It may enhance progress in the sojourners’ linguistic and general communicative abilities in that language. It may spur their intercultural awareness. It may also fulfil the ultimate educational goal of preparing for employability in the international arena. In spite of differences across countries in the nature of their respective SA programmes, SA objectives by and large are either educational, professional, or both at the same time.

Teaching Tip

Organize exchange experiences through trips abroad for all educational levels. In this way, language and intercultural sensitivity will develop, and so will individuals as a whole. Prepare learners to gain an international stance, which, as teachers, we should also have.

Following this introductory background to SA status and research, in this chapter I focus on the main issues currently investigated within the SA field of enquiry, and their relevance for the study of ISLA. After that I present a picture of the existing empirical research measuring SA effects and how individual differences may condition them. I then go on to discuss how programme features impinge on such effects. The chapter closes with a brief mention of future research directions.

Current Issues

In this section I delineate three main questions that have inspired research on the linguistic and nonlinguistic effects of SA over the years, and discuss them in relation to FI, that is, conventional language lessons taught within an educational institution, as described earlier. First, whether the common belief in the more positive effects of SA versus at-home FI is anything other than a myth. Second, if it is not, whether benefits accrue to the same extent for all SA students alike, and for all abilities. Third, in case that benefits do accrue, on what SLA theoretical grounds can they be explained.

SA Beneficial Effects, Myth or Reality?

Turning to the first issue, interestingly, common beliefs question FI for being less successful than expected. In contrast, as already presented, SA has been assumed to provide “the best opportunities to learn,” in a setting in which, as Sanz (2014) has vividly put it, according to folk beliefs “learners are immersed, soaked in the language, and feel like sponges, [. . .] They learn by doing, by living, until one day they discover themselves thinking in the language, and the ultimate experience: they dream in the language” (p. 1). Against such beliefs, Sanz claims, the existing research findings paint a less optimistic picture, as results often are either mixed or inconclusive. Hence the answer is no, SA does not always result in greater success than classroom instruction, although research also shows that, within the range of variation in results already mentioned, some learners do manage to make significant linguistic progress while abroad, in spite of others not making much (Collentine, 2009; DeKeyser, 2007; Llanes, 2011; Sanz, 2014).

The view that SA does not guarantee greater success than FI is engrained in the a priori sociolinguistic description of SA learning contexts, in contrast with FI, as presented earlier.

It has been suggested that although such a sociolinguistic description of SA as a ‘naturalistic learning context’ in contrast with at-home FI would seem to be undisputed, it may be more ‘assumed’ than real (Collentine, 2009). This, I would contend, has two explanations, of a methodological and an empirical nature, respectively. On methodological grounds, one key cornerstone in SA research is how to capture the nature of actual (variation in) the input learners receive while abroad, both in quantity and quality,

and the opportunities for interaction, with instruments other than self-reports, whose reliability is low. The widely used and adapted ‘Language Contact Profile’ (Freed, Dewey, Segalowitz, & Halter, 2004) is a fine attempt to do so; however, it does not lend itself to providing an exact quantification of ‘time on task’ with the TL while students are abroad, and still relies on self-reports. The alternative is not clear, as Devlin (2014) has emphasised,

The researcher cannot feasibly follow the learner everywhere and monitor and record all exchanges. The alternative of having a learner constantly “miked up” would prove prohibitively expensive and it is doubtful that any learner would agree. In such circumstances, researchers must rely on self-reported data.

p. 25

In recent years, a further step has been taken in this direction with the analysis of the social networks students establish while abroad. It has provided new tools in the attempt to track input and interaction patterns, albeit still relying on self-reporting (see Coleman, 2015; Dewey, 2008; Isabelli-García, 2006, among others). Informants are asked to meticulously document the frequency of interaction with their daily networks through contact diaries. Additional questionnaires enquire about further details and common contacts. The information is quantified and visualized via plot diagrams. The concept has been borrowed from sociolinguistics (Milroy, 1987) and from the rationale behind communities of practice (i.e., learning understood as the result of a social activity/practice).

On empirical grounds, the argument against the idealistic picture of SA relates to the discussion of the alleged difficulties in relating to local speakers of the TL while abroad, hence in accessing the existing ‘ideal opportunities.’ In the face of such difficulties, in order to prepare learners for SA, in at-home instruction we may need to concentrate on developing the ‘self-regulating’ capacity learners are able to display, prior to their departure abroad, that is, the extent of their proactiveness in accessing such opportunities (Dörnyei & Ushioda, 2009). This is the second issue discussed in this section, to which now we turn, that is, whether benefits accrue to the same extent for all SA students alike, and for all abilities.

SA Effects: Individual Variation in the Outcomes

In an effort to find a comprehensive operationalization of SA contexts of acquisition encompassing individual variation, a conceptual framework comprising three dimensions has been put forward within the Study Abroad and Language Acquisition (SALA) Project (Pérez-Vidal, 2014b).² It draws from the identification of ‘context’ as a key construct in research (Freed, 1995a), and ‘contact’ with the TL as a further construct put forward in a seminal publication a decade later (Collentine & Freed, 2004b). The framework allows us to relate, in a very simple manner, SA context features, learner differences, and programme design, through three dimensions.

The first dimension is represented by the *macro-level features*, that is, the sociolinguistic aspects of SA, including the amount and quality of input, interaction entailing negotiation of meaning, and output opportunities offered to learners, already discussed. The second dimension is represented by the *micro-level features*, or individual variables learners take with them when embarking on a sojourn abroad. Finally, the third dimension includes the *programme features* as revealed by the architecture of the programme, that is,

length of stay, living conditions, employment opportunities, onset language level, predeparture preparation, point in the curriculum, home academic assignments, and debriefing upon return.

Key Concept

Study Abroad can be described by means of three sets of features: The macro-level features, that is, the context features; the micro-level features, that is, what learners carry with them when they embark on a sojourn abroad; and the programme features, that is, the architecture of the specific programme they enrol on.

Teaching Tip

Programmes differ, and so will their effects. Additionally, such effects will interact with the students' own profiles, as individuals, and as language learners, and both will condition progress of any kind made while abroad.

Regarding the micro-level features of SA, it has been contended that the attributes individual students bring with them to the sojourn experience, linguistic and nonlinguistic, may help to make sense of the variation in SA proficiency outcomes, and the pervading mixed results that research has yielded thus far (Collentine, 2009; DeKeyser, 2007; Freed, Segalowitz, & Dewey, 2004; Pellegrino, 2005, to name but a few), together with the nature of the programme, as discussed further later. The individual differences that have been investigated in relation to SA include on the one hand those factors conventionally examined in the SLA literature, such as age, aptitude, attitude and motivation, gender, ethnic group and sociocultural status, personality and cognitive style, and foreign language anxiety (FLA); and on the other, those more recently investigated, the most central ones being identity, intercultural awareness, willingness-to-communicate (WTC), tolerance to ambiguity, and emotional intelligence.

Research on the interaction between such an array of differences and SA effects are beginning to throw light on the issue of 'contact' with the TL. The term 'contact' has been instrumentally used as an umbrella construct. It refers to actual access to the TL, and to opportunities for input, output, and interaction. It is the felicitous result of the self-regulatory abilities displayed by learners while abroad, to benefit from such opportunities. However, the identification of each of the constructs examined, and their operationalization, seems to be in need of further refinement, perhaps with the exception of the age factor. A summary of the existing research findings may be as follows.

Regarding the age factor, considered as the biological factor, most research has concentrated on adult learners' sojourns abroad, rather than on younger learners, actually reflecting current SA programme figures according to age. The extant research on younger learners has shown SA to be more beneficial for children (aged 10–11) than for adults (aged 19–31) in relative, not in absolute, gains (Llanes & Muñoz, 2012). Additionally, when comparing an at-home group to a SA group of young learners, a 2-month SA seems to benefit 11-year-olds in fluency, accuracy, and complexity significantly more than at-home instruction (Llanes, 2012).

Concerning aptitude, the cognitive variable, studies tapping on it, or one of its components, are definitely scarce, perhaps due to the experimental nature of its methods and

tests. Sunderman and Kroll (2009) investigated the relationship between internal cognitive resources, in the form of working memory resources in lexical comprehension and production and benefits while abroad, finding that reaching a certain threshold of working memory resources is a necessary condition to benefit from the study-abroad context in terms of accurate L2 production, along the lines of Collentine's (2009) threshold level discussed earlier. O'Brien, Segalowitz, Freed, and Collentine (2007) found that phonological memory, another of the components of aptitude, predicted oral gains while abroad. Segalowitz and Freed (2004) analysed general oral proficiency and oral fluency, and the relation these oral gains bore to (1) reported hours spent in extracurricular language activities and (2) L2-specific cognitive measures of speed of lexical access (word recognition), efficiency (automaticity), and speed and efficiency of attention control hypothesized to underlie oral performance. Similarly, Taguchi (2008) investigated how processing speed and contact hours correlated with the ability to comprehend pragmatic intentions in the case of Japanese learners of English, finding that lexical access speed and contact hours significantly correlated with comprehension, but not accuracy. Still within the general cognitive processing abilities involved in the aptitude construct, lexical access and attention control seem to condition gains made in both SA and at-home learning contexts (Kormos & Sáfár, 2008; Segalowitz & Freed, 2004).

If we now turn to the emotional or affective variables, regarding attitude and motivation, linguistic self-confidence and intended effort has shown to rise significantly during FI, while desiring to live in a different country from one's own is higher after SA; as are foreseeing better career prospects, wanting to travel, and wishing to meet new people (Juan-Garau & Trenchs-Parera, 2014). As for personality, Dewaele, Comanaru, and Faraco (2015) have established that "identity is more fluid, socially constructed and constrained, and contextually determined, whereas personality [. . .] is generally thought to be a more stable construct" (p. 109) including five traits. On the basis of the Multicultural Personality Questionnaire, and a reflective interview taking place at the end of a sojourn abroad, the authors found that 77% of the participants had changed in one trait, namely emotional stability, and they felt more confident, resourceful, and autonomous upon return. Identity, as a less stable trait, has been established in a complex dynamic interplay between individual agency, biology, and societal imposition. Individuals can deploy a number of identities (Dörnyei & Ushioda, 2009; Pavlenko & Blackledge, 2004; Regan, 2010). Identity embraces gender, race, ethnicity, class, sexuality, and urbanity (Devlin, 2014). As already stressed, the need to and difficulty of repositioning identities in an L2 culture can influence the contact learners have with native speakers (Kinging, 2009; Pavlenko & Piller, 2001).

Regarding intercultural sensitivity, the Development Model of Intercultural Sensitivity (DMIS) (Bennet, 1986), and the Intercultural Development Inventory (IDI) (Hammer, Bennet, & Wiseman, 2003) have been proposed to measure intercultural awareness, and used primarily with individuals who must start functioning in international settings. These two instruments have been developed within the paradigm that defines culture as the knowledge, motivation, and skills needed to interact effectively and appropriately with members of different backgrounds (Byram, 1997). Cultural sensitivity is key to language acquisition, as intercultural contact has been highlighted as being influential on language learning motivation, which in turn has been found to be directly related to acquisition (Dörnyei & Csizér, 2005; Kormos, Csizér, & Iwaniec, 2014). Hismanoglu (2011) found higher proficiency students had higher intercultural awareness, which develops further with pre-training. Willingness to communicate (WTC), that is the capacity to decide to engage in L2 interactions (Dewaele, 2007) and foreign language anxiety (FLA), or the worry and

usually negative emotional reaction arising when learning or using an L2 (MacIntyre, 2007), have also been investigated showing to be positively influenced by a SA, so that more WTC and less anxiety seem to obtain during SA (see Dewaele & Wei, 2013; Dewaele et al., 2015, respectively).

Key Concept

Managing access to TL speakers: Learners abroad may have plenty of opportunities to communicate with TL speakers, and also be exposed to the media and the linguistic landscape. This contact should allow them to receive massive amounts of meaningful input while interacting and producing output, both conducive to language acquisition. However, not all learners avail themselves of such opportunities. They may lack the abilities necessary to establish contact with TL speakers, which are dependent on their age, aptitude, motivation and attitudes, affects such as foreign language anxiety (FLA), and willingness to communicate (WTC). Learner awareness of other cultures and identity are also important.

Teaching Tip

SA sojourners must be equipped and be able to deploy an array of strategies in order to establish contact with TL speakers. This should allow them to benefit from input, output and interaction practice in the different settings, registers, channels, topics, and degrees of formality encountered while abroad, in both in-class and out-of-class situations.

While waiting for further studies to confirm the extant research findings on the impact of individual differences, the theoretical and empirical basis attesting to the linguistic benefits of SA, to which I now turn, stands on somewhat firmer ground.

SA Beneficial Effects: Bridging the Gap Between SA and SLA Theory

The third and last issue discussed in this section relates to the efficacy of SA and the basis on which it can be explained. This relationship may be approached in terms of the psycholinguistic mechanisms that come into play in the SA naturalistic context. In order to support the argument that SA is an optimal context for language development, some authors have invoked classic SLA theories that adopt an interactionist framework to describe (comprehensible) input, interaction with containing negotiation of meaning, and output as the necessary conditions for acquisition (see for example Sanz, 2014). Indeed, SA may offer plenty of opportunities to negotiate meaning, which interactionists consider to be the locus of language acquisition. From a cognitive processing perspective, SA provides opportunities for implicit language learning, as opposed to the explicit attention to form and rules typical of at-home classroom instruction contexts, as described earlier. However, something is often missing in SA, which FI does offer; DeKeyser (2007) argues that opportunities for feedback, considered to be important

for linguistic development, are missing. All in all, if we assume that SA contexts are superior to at-home classroom instruction, the question, or the problem, is where does that leave classroom instruction? First of all, it must be borne in mind that implicit approaches to language learning also obtain in FI contexts, as communicative language teaching and the popular task-based approaches illustrate, and, vice versa, attention to form is also present during SA; as Hassall (2013) points, in his study on SA pragmatics, or as Sanz (2014) has noted, by referring to “Schmidt’s seminal work on attention in SLA, which started with observations of what he labelled *noticing* in the diary he kept while living abroad in Brazil” (p. 2). In consequence, the gap between the two types of practice, formal and naturalistic is somewhat blurred.

Interestingly, one possible answer to the previous question may come from skill acquisition approaches to SA as ‘foreign language practice,’ which allows us to bridge the gap between SA and at-home classroom instruction. In a nutshell, from the perspective of skill acquisition theory, three stages can be distinguished in terms of the practice needed for language learning: (1) declarative knowledge, that is explicit knowledge of rules; (2) proceduralization, that is the process of coming to terms with rules for future conscious retrieval; and (3) automatization. SA is most conducive to automatization, because it can provide “the large amount of practice necessary for the gradual reduction of reaction time, error rate, and minimal interference with other tasks that characterize the automatization process” (DeKeyser, 2007, p. 216). Classroom instruction is where declarative knowledge is established, and, in turn, the first stages of proceduralized knowledge occur, that is, practice with conscious use of rules. Then, what students can do while they are abroad is to proceed to the further stages of proceduralization, whereby the process of conscious retrieval and use of rules is speeded up, leading them naturally toward automatization. While doing that, learners put to play all the abilities, skills and strategies learnt in the classroom for the purpose of communicating in natural circumstances. Such views have generated a working hypothesis, which states that it is the combination of at-home classroom instruction and SA that yields the largest proficiency gains in learners (Pérez-Vidal, 2014b). In what follows, a description of the main research findings in SA concerning such gains is presented.

Key Concept

SA contexts allow students to make use of opportunities to practice the language, in a way that complements what they have been practicing in the classroom, particularly at home prior to the sojourn, and also during their sojourn, thus speeding up all linguistic progress.

Teaching Tips

Learners must be able to bridge the gap between FI knowledge and strategies developed at home, on the one hand, and the use of this knowledge and these strategies during their sojourns abroad.

Empirical Evidence

This section presents research findings on the linguistic effects of SA sojourns spent in TL destinations in greater detail than previously. The conventional language abilities and skills are reviewed. Short and mid-term effects are considered and a note on the methods used is made.

SA Learning Contexts: Disentangling the Linguistic Effects

As mentioned, SA research in the 1990s became a prolific enterprise that gave way to publications including both holistic multidimensional studies, focusing on skill development, and studies focusing on discrete linguistic areas. Some were conducted in the US, such as the analysis of 668 students by Brecht, Davidson, and Ginsberg (1995), who found SA to be one of the strongest predictors of development in reading, speaking, and listening in L2 Russian; others investigated European students, such as Milton and Meara's (1995) study on vocabulary, which found that of the 586 sojourners in the UK, the lower-level learners were the ones with the most improvement in their lexical knowledge. More focused studies followed, such as those looking into oral production. Towell, Hawkins, and Bazergui (1996) identified speech rate and mean length of run, that is the average number of syllables produced between pauses, as the main components of improvement in UK students' L2 French fluency. However, DeKeyser's (1991) analysis of US students' grammatical and vocabulary development together with oral proficiency, found no differential effect between SA and at-home instruction.

More recently, Collentine and Freed's (2004a) monograph pushed the linguistic scope of research to focus more decidedly on the variables conducive to gains, while presenting the 'contact' factor discussed earlier as exerting influence on outcomes. Collentine (2009) later summarized: "Interestingly, linguistic aspects that do indeed seem to benefit from SA, such as fluency and discursive abilities, are often not those in which at home foreign language program directors hope to see improvements, such as grammatical aspects" (p. 222). Indeed, results for grammar and linguistic complexity have been mixed (DeKeyser, 1991; Howard, 2005). Phonological development, an area with no more than a handful of published studies, has not been seen to improve to a larger degree while abroad than at home (Díaz-Campos, 2004), or even shows higher improvement in FI at home (Mora, 2014).

Turning to the studies tapping into skill development, listening skills have been reported to increase significantly more during SA than in FI (Kington, 2009; Llanes, 2011). Reading has also been shown to improve during SA (Dewey, 2004). Regarding written skills, for which again no more than a handful of studies exist, Sasaki (2011) found that lasting effects of SA on participants' writing were seen to be determined by whether they were able to create 'imagined communities,' an issue of identity. In addition, sociolinguistic aspects of language use also appear to develop substantially (Regan, Howard, & Lemée, 2009), and so do pragmatic abilities, in particular those associated with the use of formulaic routines, as part and parcel of fluency (see Dewaele & Regan, 2001; Pérez-Vidal & Juan-Garau, 2009, and more recently Imura & Shimizu, 2012). Hallal (2013) is a case in point in showing how, for a group of Australian students, even short stays were beneficial for acquiring terms of address in Indonesian. Oral proficiency, and especially fluency, is the area showing most benefits after SA, even for short stays of 3–4 weeks (Llanes & Muñoz, 2009). However, negative aspects have also been revealed, such as those shown in Hallal (2013): how transfer

from the L1 and instruction can hinder development, and also being identified by the locals as a ‘foreigner.’

Regarding the nature of the empirical studies reviewed, the approach taken by most has been to measure students with a pretest before departure and posttest upon return, sometimes adding a ‘while-away’ test. In recent years, on the wake of methodological reviews pointing to the lack of control conditions in such research (see for example Rees & Klapper, 2008), studies have incorporated FI in the home institution, or immersion in the classroom, as control conditions. The latter is the case of Freed, Segalowitz et al. (2004) who compared SA with at-home FI and with domestic immersion, only to find that domestic immersion yielded the highest benefits on oral fluency, a fact they attributed to students in domestic immersion investing larger amounts of out-of-class time on task (around 3 hours per day) than those in FI. It may be concluded that at such advanced levels, progress comes not with simply interacting with TL speakers, but with engaging in more cognitively demanding out-of-class activities, such as academic work. Serrano, Llanes, and Tragant (2011) compared written and oral performance following intensive and semi-intensive domestic FI on the one hand, and SA on the other. They found that there were differences between gains acquired abroad, and gains acquired as a result of semi-intensive courses at home, but not with intensive courses at home, which proved to be as beneficial as SA, confirming the positive effects of ‘domestic immersion’ as in Freed, Segalowitz et al. (2004).

Finally, few studies have taken up the issue of the long-term effects of SA, as noticed by Llanes (2011). A recently published volume deserves attention on this and other issues: Pérez-Vidal (2014a) reports on the results of a multimeasures, mixed methods compilation of 10 studies from the SALA project. It examines the benefits, short and mid-term, of a 3-month ERASMUS exchange in an English-speaking country, following a 6-month period of FI at home (Pérez-Vidal, 2014a). Participants were a homogeneous group of 80 advanced-level students. They were tested with a repeated-measures design, after experiencing each context, and with a last data collection tapping into retention effects taking place, 15 months upon return. Learners were used as their own matched pair, in that, following Milton and Meara (1995), their proficiency at the end of the FI period was contrasted with that at the end of the SA period, in a within-groups design. Results showed that the progress learners made as a result of SA was superior to that made in the FI context, “in oral skills as measured through integrative tasks, as far as fluency and accuracy are concerned, and also listening, but not in phonological development regarding both production and perception; results for the latter are even significantly better at home” (Mora, 2014, p. 189). Writing and lexico-grammatical abilities also improved significantly. Positive effects were maintained in the long run. Motivation and beliefs differed whether at home or abroad, and intercultural awareness significantly improved while abroad, but gains were not maintained and students returned to pretest levels. Pérez-Vidal’s (2015b) further analysis of the ‘relative’ gains made abroad by the same learners showed the most benefit for oral skills, both receptive and productive, except for phonology, and also for written skills, except fluency and lexico-grammatical ability.

In sum, the findings of the existing research point at SA favourably impacting on pragmatics, writing, oral production, and reception, with the exception of phonological development, which would appear to benefit most from a FI context, where attention to form often prevails over attention to meaning. Interestingly, however, the linguistic impact of domestic immersion with plenty of out-of-class practice seems to be equally beneficial to SA, something that would confirm the idea that ‘time on task’ is what matters. Concerning levels of accuracy and lexico-grammatical abilities, recent studies seem to further prove that the benefits of the

massive amounts of practice SA allows students to have, seemingly lead to automatization, certainly so in the case of advanced-level learners on European exchanges (as DeKeyser, 2014 has stressed, but see the discussion on programme features later in this chapter). This point takes us to the next section, which deals with programmatic considerations.

Pedagogical Implications

SA programmes vary to a large extent, hence it must be expected that their effects will, in turn, also vary, in a similar way as learners' differences yield variation in SA effects, as already presented. In this section a discussion is offered on the impact on language acquisition of SA programme features and their interactions. They are grouped according to the classification into the aforementioned eight features (Pérez-Vidal, 2014b). It cannot be over-emphasized that care in the design and implementation of SA programmes can only result in their greater efficacy and outcomes. The first four features include the philosophy of the programme, length of stay, housing arrangement and onset language level.

The Philosophy of the Programme

Programmes and their design may differ across the world in their objectives. In different programme designs, there is to be found a binary tension, with full integration and acculturation in the community at one end (for instance the ERASMUS programmes, see Beattie, 2014), and lesser degrees of integration at the other end (the case of the sheltered/island-programmes in which students travel with instructors and do not attend courses with local students, see Kinginger, 2009).

Key Concept

SA programmes: SA programmes should be accountable in terms of their efficacy and success.

Teaching Tip

Adequately prepare administrative professionals and language instructors on exchange programmes to better guide students in the practical and academic matters coming into play for the success of exchange programmes, before, during, and after the programme.

Length of the Programme

The question of 'how long is long enough' in SA programmes has been underinvestigated (DuFon & Churchill, 2006). Most empirical research seems to prove that the longer the better, such as Sasaki's (2011) seminal study on writing. Dewaele et al. (2015) found that stays longer than 1 year have a larger effect on WTC. One year was better than a semester for the development of reading and writing skills of L1 English learners of German (Fraser, 2002). Ten months were better than 3 months in Hoffman-Hicks's (2000) study on pragmatic

Carmen Pérez-Vidal

abilities in L2 French, and 9 months better than 6 months in Ife, Vives-Boix, and Meara's (2000) study on vocabulary. All in all, however, there is evidence that 2-month programmes (Llanes, 2012) and programmes of less than 1 month can already yield significant gains in oral fluency, accuracy, and listening (Llanes & Muñoz, 2009).

DeKeyser (2014) posits a clarifying relationship between onset level and length of stay “automatizing a limited number [low level] of highly frequent elements leads to more gains in the short run, while students with a much larger number of elements to automatize [higher level] only outpace the less advanced ones in the long run” (p. 317), something which Lara, Mora, and Pérez-Vidal's (2014) study confirms.

Housing Arrangement

According to Kinginger (2014), no association has been found between the SA living arrangements (halls, college dorms, shared flats, or family settings) and the development of proficiency. She further points at an age issue:

younger learners may be more likely than their more aged peers to be received *in loco parentis* as temporary children, and to tolerate and benefit from this arrangement more easily. [. . .] [and report] numerous opportunities to interact in various settings involving all generations of their host families and the families' social networks.

Kinger, 2014, p. 54

However, Wilkinson (1998) discloses limited interaction in home stays.

Onset Language Level

Onset language level has been found to be a clear predictor of linguistic gains while abroad, in relation to length of stay as discussed earlier, and in general. DeKeyser (2007) refers to a *functional level*, equivalent to an intermediate-advanced level, which should allow learners to complete the proceduralization process and begin with automatization, while engaging in communicative interaction. Collentine (2009) believes that “there is a *threshold level* which learners must reach to benefit fully from the SA context of learning: There are most likely specific domains that require a particular developmental threshold for specific gains to occur” (p. 221). However, Llanes and Muñoz's (2009) participants with lower proficiency level made comparatively greater gains after a 3–4 week stay “in using L2 words (. . .) and in producing more accurate and fluent speech” (p. 10), and so did Pérez-Vidal's (2014a) university participants in the SALA studies presented.

A final set of four features rounds up the architecture of SA programmes: their academic dimension, the predeparture preparation, the point in the curriculum, and debriefing upon return.

The Academic Dimension of the Programme

The academic dimension of the programme relates to the quantity and quality of both classroom language and out-of-class language practice required from students. It has been found to bear an impact on linguistic progress. Segalowitz and Freed (2004), found a relationship between the previously discussed feature of onset language level to academic work, found that “initial oral performance levels may also influence learners' predispositions to

extracurricular communicative opportunities (listening to radio, films, and television) [. . .], oral fluidity correlating with reported extracurricular reading” (p. 195).

Predeparture Preparation

Efforts should be geared toward predeparture preparation to help learners maximize opportunities for interaction, deal with fears, affects, and intercultural development while abroad (Chieffo & Zipser, 2001; Collentine, 2009). An example of such kind of preparation with a practical orientation are the *Study Abroad Self-Study Guides* aimed at students, programme professionals, and language instructors developed and subsequently tested by Paige et al. (2004). Their findings point at strategy training prior to and while abroad significantly improving strategy use when speaking and listening, and learning culture.³

Key Concept

Predeparture preparation: Being well briefed for SA is a necessary provision to allow more students than those who are genuinely talented, to benefit from a sojourn abroad; it should include guidance to develop learner autonomy and strategies to improve linguistic, intercultural, and communicative skills, to be used when in the host country.

Teaching Tip

Offer a Preparation Module for students to follow before embarking on an exchange, or use readily available materials (see Paige et al., 2004; see also a recently published web www.intClass.org). It should train them to develop strategies to benefit from the interaction opportunities while abroad. Diary writing has been recommended as a useful tool to help students develop as autonomous students, set themselves linguistic and cultural objectives, and gain linguistic and intercultural awareness while abroad. Assignments that force students to interact with members of the community are beneficial.

Stage in the Curriculum

The SA experience is recommendable for both secondary school and university level students. Following DeKeyser (2007), from the perspective of a skill acquisition theory “the transition in skill acquisition that *should* coincide with going abroad is automatization. As this process requires a very large amount of varied practice, the native-speaking environment is a much better context for it than the classroom” (p. 217, emphasis in original). The impact of learners’ stage of development on linguistic progress while abroad seems thus evident.

Debriefing Upon Return

Students come back from SA having lived a unique experience. Debriefing should ideally be built-in in the exchange programme to capture such a momentum, be able to keep the friendships they established abroad, particularly when the TL was their medium of communication, and reap the benefits once back (adequate courses, exam certifications).

Key Concept

Debriefing upon return: One of the most neglected aspects of SA programmes, which deserves attention to catch the momentum of the advances made by students abroad, is debriefing.

Teaching Tip

Satisfaction questionnaires may be used for debriefing, discussing quantity and quality of work done, showing lecture notes and assignments done abroad. Alternatively, diary keeping can tap into learning strategies, language awareness, intercultural encounters, and so forth. Open interviews with focus groups can be organized. SA experienced students can help orientate the following cohort going abroad. Returnees should be encouraged to keep all contacts made abroad, using their TL, to take relevant FI courses, and eventually aim at formally certifying their newly acquired linguistic and intercultural skills.

In sum, the puzzle between the two seemingly antagonist contexts of learning here discussed, that is, SA and FI, can be solved when we see them along yet another “continuum of practice from basic classroom instruction to pre-departure training, on-site observation and guidance, and courses for students returning home” (DeKeyser, 2007, p. 208).

Key Concept

SA programme design: SA programmes can be seen to differ around eight variables, which reflect their core features: philosophy of the programme, length, accommodation arrangements, onset level, academic work while abroad, predeparture preparation, point in the curriculum, and debriefing upon return.

Teaching Tips

- Short stays can prove significantly effective, but longer stays are even more effective, particularly for advanced learners who need more time to show progress.
- Home accommodation with families may be very fruitful, particularly with younger learners; dorms may be better for adult learners.
- Advanced level learners may need more demanding activities organized for them, as they seem to be at ceiling in contrast with lower-level learners who always make significantly greater progress.

Future Directions

To close this chapter, a brief presentation of future directions for the SA research agenda is offered, as suggested by a number of authors, often following critiques of the extant research, mostly on methodological grounds (Collentine, 2009; DeKeyser, 2014; Llanes, 2011; Rees & Klapper, 2008; Sanz, 2014, to name a few). Problems may be identified with respect to the following issues. Language level testing is often organized only after the period spent abroad, and this with tests that may not have been piloted, and for which validity and reliability measures are not given. As for the nature of the tests used to quantify results, both broad and fine-grained tests are seldom used, neither are in-depth questionnaires with an inductive–deductive approach; statistical analyses are not always incorporated. As already discussed, we lack an instrument to systematically account for input and interaction features while abroad, and to conduct objective observations, synchronic and diachronic throughout the sojourn. Finally, aptitude and working memory are often not controlled, and programme features not used as independent variables.

All in all, it would seem that the practice of using mixed methods is still not common, although efforts have already been made in such a direction. Greater rigour and control is also necessary (Sanz, 2014). DeKeyser (2014) sets very clear goals for research on some of the issues that remain unresolved. We need to examine the interaction between individual features and the type of practice indulged in within different environments during SA, while using observation and learners' introspective protocols to elicit learners' attitudes, beliefs, motivations, and emotions, and profiling features in interlocutors. To the previous elements, laboratory methods geared toward focusing on psycholinguistic processes involved in skills development on specific linguistic features should be added, such as in the case of less salient structures or abstract vocabulary, for example, within situations with extensive practice. In sum, the research agenda is extensive.

I would like to close this chapter with the voice of the students, who beyond such research challenges most surely experience SA above all as a life-changing event of huge personal growth. This is illustrated in the following fragment taken from a debriefing text written by a Spanish university student returning from a 3-month SA sojourn:

Now it's been a few weeks since I returned to Spain from England. Looking back, I see myself before the stay abroad very different from who I am now. This experience has taught me much more than what I could've ever imagined and defining it as amazing falls short. Not only have I acquired some new habits—drinking a lot of tea, and paying a great deal of attention to the weather, for instance—but, most importantly, I am definitely much more open-minded and somehow more mature and responsible. The people, the place and the memories are unquestionably the best I've taken from the experience and will always remain in my mind.

Notes

1. It must be noted that SA may also involve internships (see Devlin, 2014; Mitchell, McManus, and Tracy-Ventura, 2015) on work placements (Tracy-Ventura et al., 2016).
2. The Study Abroad and Language Acquisition (SALA) project is a long-standing research project, currently funded by the Ministry of Economy and Competitiveness, based in Barcelona and Palma de Majorca, in the Balearic Islands, both in Spain (FFI2013–48640.C2–1-P), and the Generalitat de Catalunya (2014 SGR 1586), respectively.

Carmen Pérez-Vidal

3. Paige et al. (2004) defined the aim of the guides in the following terms:

[they] would (1) be generalizable across study abroad sites, cultures, and languages, (2) emphasize a strategies-based approach to language and culture learning, (3) address all three phases of the experience (pre-departure, in-country, and re-entry), (4) assist students, program professionals, and language instructors, (5) be based on theory and research about language acquisition and intercultural competence, and (6) be flexible in their application—they could be used in a self-study format (Students' Guide), an orientation program, and a formal course. (p. 255)

References

- Backhaus, P. (2007). *Linguistic landscapes: A comparative study of urban multilingualism in Tokyo*. Clevedon: Multilingual Matters.
- Banks, M., & Bhandari, R. (2012). Global student mobility. In D. K. Deardorff, H. DeWit, J. Heyl, & T. Adams (Eds.), *The Sage handbook of international education* (pp. 379–398). Thousand Oaks, CA/London: Sage.
- Beattie, J. (2014). The “ins and outs” of a study abroad programme: The SALA exchange programme. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 59–87). Amsterdam/Philadelphia: John Benjamins.
- Bennet, M. J. (1986). A developmental approach to training for intercultural sensitivity. *Intercultural Journal for International Relations*, 10, 179–196.
- Block, D. (2003). *The social turn in second language acquisition*. Edinburgh: Edinburgh University Press.
- Brecht, R., Davidson, D., & Ginsberg, R. (1995). Predictors of foreign language gain during study abroad. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37–65). Amsterdam/Philadelphia: John Benjamins.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon: Multilingual Matters.
- Chieffo, L., & Zipser, R. (2001). Integrating study abroad into the foreign language curriculum. *ADFL Bulletin*, 32, 79–85.
- Coleman, J. (1998). Language learning and study abroad: the European perspective. *Frontiers*, 4, 167–203.
- Coleman, J. (2002). Translating research into disseminated good practice: The case of student residence abroad. In R. McDonald & J. Wisdom (Eds.), *Academic and educational development: Research, evaluation and changing practice in higher education* (pp. 88–98). London: Kogan Page.
- Coleman, J. (2006). English-medium teaching in European higher education. *Language Teaching*, 39(1), 1–14.
- Coleman, J. (2015). Social circles during residence abroad: What students do, and who with. In R. Mitchell, N. Tracy-Ventura, & K. McManus (Eds.), *Social interaction, identity and language learning during residence abroad* (Eurosla Monograph Series 5, pp. 33–50). Amsterdam: The European Second Language Association.
- Collentine, J. (2009). Study abroad research: Findings, implications, and future directions. In M. H. Long & C. Doughty (Eds.), *The handbook of second language teaching* (pp. 218–233). Malden, MA: Blackwell.
- Collentine, J., & Freed, B. (2004a). *Studies in Second Language Acquisition*, 26(2), 153–172.
- Collentine, J., & Freed, B. (2004b). Introduction: Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition*, 26(2), 153–171.
- DeKeyser, R. (1991). Foreign language development during a semester abroad. In B. Freed (Ed.), *Foreign language acquisition: Research and the classroom* (pp. 104–119). Lexington, MA: D. C. Heath.
- DeKeyser, R. (2007). Study abroad as foreign language practice. In R. DeKeyser (Ed.), *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 208–226). Cambridge: Cambridge University Press.

- DeKeyser, R. (2014). Research on language development during study abroad: Methodological considerations and future perspectives. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 313–327). Amsterdam/Philadelphia: John Benjamins.
- Devlin, A. M. (2014). *The impact of study abroad on the acquisition of sociopragmatic variation patterns: The case of non-native speaker English teachers*. Bern: Peter Lang.
- Dewaele, J. M. (2007). The effect of multilingualism, sociobiographical and situational factors on communicative anxiety and foreign language anxiety of mature language learners. *International Journal of Bilingualism*, 6, 23–3.
- Dewaele, J. M., Comanaru, R. S., & Faraco, M. (2015). The affective benefits of a pre-session course at the start of study abroad. In R. Mitchell, N. Tracy-Ventura, & K. McManus (Eds.), *Social interaction, identity and language learning during residence abroad* (Eurosla Monograph Series 5, pp. 33–50). Amsterdam: The European Second Language Association.
- Dewaele, J. M., & Li, Wei (2013). Is multilingualism linked to a higher tolerance of ambiguity? *Bilingualism: Language & Cognition*, 16(1), 231.
- Dewaele, J. M., & Regan, V. (2001). The use of colloquial words in French advanced interlanguage. *EUROSLA Yearbook*, 1, 51–68.
- Dewey, D. (2004). A comparison of reading development by learners of Japanese in intensive domestic immersion and study abroad contexts. *Studies in Second Language Acquisition*, 26(2), 303–327.
- Dewey, D. (2008). Japanese vocabulary acquisition by learners in three contexts. *Frontiers: The Interdisciplinary Journal of Study Abroad (Fall)*, 127–148.
- DeWit, H., & Merckx, G. (2012). The history of internationalisation of higher education. In D. K. Deardorff, H. De Wit, J. Heyl, & T. Adams (Eds.), *The Sage handbook of international education* (pp. 43–60). Thousand Oaks, CA/London: Sage.
- Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, 26(2), 249–273.
- Dörnyei, Z., & Csizér, K. (2005). The effects of intercultural contact and tourism on language attitudes and language learning motivation. *Journal of Language and Social Psychology*, 24(4), 327–357.
- Dörnyei, Z., & Ushioda, E. (2009). *Motivation, language identity and the L2 self*. Clevedon: Multilingual Matters.
- DuFon, M. A., & Churchill, E. (Eds.). (2006). *Language learners in study abroad contexts*. Clevedon: Multilingual Matters.
- European Commission. (1995). *Teaching and learning: Towards a learning society: 449 white paper on education and learning*. Brussels: DGV.
- European Commission. (2013). *On the way to ERASMUS+: A statistical overview of the ERASMUS programme in 2011–12*. Retrieved from <http://ec.europa.eu/education/>
- European Commission. (2016). New findings about mobility: International experience is valued in Eastern Europe, mobile Southern European students move abroad for jobs and students under 23 benefit more from mobility (Press release).
- Falk, R., & Kanach, N. (2000). Globalization and study abroad: An illusion of paradox. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 6, 155–168.
- Fraser, C. C. (2002). Study abroad: An attempt to measure the gains. *German as a Foreign Language Journal*, 1, 45–65.
- Freed, B. F. (Ed.). (1995a). *Second language acquisition in a study abroad context*. Amsterdam/Philadelphia: John Benjamins.
- Freed, B. (1995b). Language learning and study abroad. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 3–33). Amsterdam/Philadelphia: John Benjamins.
- Freed, B. F., Dewey, D., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition*, 26(2), 349–356.
- Freed, B. F., Segalowitz, N., & Dewey, D. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(2), 275–301.
- Gore, J. (2005). *Dominant beliefs and alternative voices: Discourse, belief and gender in American study abroad*. New York: Routledge.

Carmen Pérez-Vidal

- Hammer, M. R., Bennet, J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *Journal of Intercultural Relations*, 27(4), 421–443.
- Hassall, T. (2013). Pragmatic development during short-term study abroad: The case of address terms in Indonesian. *Journal of Pragmatics*, 55, 1–17.
- Hismanoglu, M. (2011). An investigation of ELT students' intercultural communicative competence in relation to linguistic proficiency, overseas experience and formal instruction. *International Journal of Intercultural Relations*, 35(6), 805–817.
- Hoffman-Hicks, S. D. (2000). The longitudinal development of French foreign language pragmatic competence: Evidence from study abroad. *Dissertations Abstract International*, 61(2), 561.
- Howard, M. (2005). Second language acquisition in a study abroad context: A comparative investigation of the effects of study abroad and formal language instruction on the L2 learner's grammatical development. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 495–530). Berlin: Mouton de Gruyter.
- Ife, A., Vives Boix, G., & Meara, P. (2000). The impact of study abroad on the vocabulary development of different proficiency groups. *Spanish Applied Linguistics*, 4(1), 55–84.
- Imura, S., & Shimizu, T. (2012). Development of formulaic language in adult second language learners' lexicon in study abroad acquisition. *Language, Culture and Communication*, 4, 173–190.
- Isabelli-García, C. L. (2006). Study abroad social networks, motivation, and attitudes: Implications for SLA. In M. DuFon & E. Churchill (Eds.), *Language learners in study abroad contexts* (pp. 231–258). Clevedon: Multilingual Matters.
- Jackson, J. (2008). *Language, identity and study abroad: Sociocultural perspectives*. London: Equinox.
- Jackson, J. (2010). *Intercultural journeys: From study to residence abroad*. Basingstoke: Palgrave MacMillan.
- Jackson, J. (2013). Pragmatic development in study abroad contexts. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Hoboken, NJ: Wiley-Blackwell.
- Juan-Garau, M., & Salazar-Noguera, J. (Eds.). (2015). *Content-based learning in multilingual educational environments*. Berlin: Springer.
- Juan-Garau, M., & Trenchs-Parera, M. (2014). A longitudinal study of learners' motivation and beliefs in at home and study abroad. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 258–281). Amsterdam/Philadelphia: John Benjamins.
- Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Oxford: Blackwell.
- Kinginger, C. (2009). *Language learning and study abroad: A critical reading of research*. Houndmills: Palgrave Macmillan.
- Kinginger, C. (2014). Language socialization in the homestay: American high school students in China. In R. Mitchell, N. Tracy-Ventura, & K. McManus (Eds.), *Social interaction, identity and language learning during residence abroad* (Eurosla Monograph Series 5, pp. 33–53). Amsterdam: The European Second Language Association.
- Kormos, J., Csizér, K., & Iwaniec, J. (2014). An interview study of intercultural contact and its role in language learning in a foreign language environment. *Journal of Multilingual and Multicultural Development*, 35(2), 51–166.
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11(2), 261–271.
- Lara, A. R., Mora, J. C., & Pérez-Vidal, C. (2014). How long is long enough? L2 English Development through study abroad programmes varying in duration. *Innovation in Language Learning and Teaching*, 9(1), 1–12.
- Leask, B. (2015). *Internationalizing the curriculum*. New York: Routledge.
- Llanes, A. (2011). The many facets of study abroad: An update of the research on L2 gains emerged during a SA experience. *International Journal of Multilingualism iFirst*, 1–27.
- Llanes, A. (2012). The short and long-term effect of a short study abroad experience. *System*, 40, 179–190.
- Llanes, A., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37, 353–365.

- Llanes, A., & Muñoz, C. (2012). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 64(1), 1–28.
- MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *Modern Language Journal*, 91, 564–576.
- Milroy, I. (1987). *Language and social networks*. Oxford: Blackwell.
- Milton, J., & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL Review of Applied Linguistics*, 107–108, 17–34.
- Mitchell, R., McManus, K., & Tracy-Ventura, M. (2015). Placement type and language learning during residence abroad. In R. Mitchell, N. Tracy-Ventura, & K. McManus (Eds.), *Social interaction, identity and language learning during residence abroad* (Eurosla Monograph Series 5, pp. 95–115). Amsterdam: The European Second Language Association.
- Mitchell, R., Tracy-Ventura, M., & McManus, K. (Eds.). (2015). *Residence abroad, social networking and second language learning*. Eurosla Monograph Series 4. Amsterdam: The European Second Language Association.
- Mora, J. C. (2014). The role of onset level on L2 perceptual phonological development after formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 167–195). Amsterdam/Philadelphia: John Benjamins.
- Murphy-Lejeune, E. (2002). *Student mobility and narrative in Europe: The new strangers*. London: Routledge.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29(4), 557–581.
- Paige, R. M., Cohen, A., & Shiveley, R. (2004). Assessing the impact of a strategies-based curriculum on language and culture learning while abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 253–276.
- Pavlenko, A., & Blackledge, A. (Eds.). (2004). *Negotiation of identities in multilingual contexts*. Clevedon: Multilingual Matters.
- Pavlenko, A., & Piller, I. (2001). Introduction. In A. Pavlenko, A. Blackledge, I. Piller, & M. Teutsch-Dwyer (Eds.), *Multilingualism, second language learning and gender* (pp. 1–16). Berlin: Mouton de Gruyter.
- Pellegrino, V. A. (2005). *Study abroad and second language use: Constructing the self*. Cambridge: Cambridge University Press.
- Pérez-Vidal, C. (Ed.). (2014a). *Language acquisition in study abroad and formal instruction contexts*. Amsterdam/Philadelphia: John Benjamins.
- Pérez-Vidal, C. (2014b). Study abroad and formal instruction contrasted: The SALA Project. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 17–59). Amsterdam/Philadelphia: John Benjamins.
- Pérez-Vidal, C. (2015a). Languages for all in education: CLIL and ICLHE at the crossroads of multilingualism, mobility and internationalization. In M. Juan-Garau & J. Salazar-Noguera (Eds.), *Content-based learning in multilingual educational environments* (pp. 31–51). Berlin: Springer.
- Pérez-Vidal, C. (2015b). Practice makes best: Contrasting learning contexts, comparing learner progress. *International Journal of Multilingualism*, 12(4), 453–470.
- Pérez-Vidal, C., & Juan-Garau, M. (2009). The effect of study abroad (SA) on written performance. *EUROSLA Yearbook*, 9, 270–296.
- Rees, J., & Klapper, J. (2008). Issues in the quantitative longitudinal measurement of second language progress in the study abroad context. In L. Ortega & H. Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 89–105). London: Routledge.
- Regan, V. (1995). The acquisition of sociolinguistic native speech norms: Effects of a year abroad on second language learners of French. In Freed, B. (Ed.), *Second language acquisition in a study abroad context* (pp. 245–268). Amsterdam/Philadelphia: John Benjamins.
- Regan, V. (2010). Sociolinguistic competence, variation patterns and identity construction in L2 and multilingual speakers. In L. Roberts, M. Howard, M. O'Laoire, & D. Singleton (Eds.), *EUROSLA Yearbook* (pp. 21–37). Amsterdam/Philadelphia: John Benjamins.

Carmen Pérez-Vidal

- Regan, V., Howard, M., & Lemée, I. (2009). *The acquisition of sociolinguistic competence in a study abroad context*. Clevedon: Multilingual Matters.
- Sanz, C. (2014). Contribution of study abroad research to our understanding of SLA processes and outcomes: The SALA project, an appraisal. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 1–17). Amsterdam/Philadelphia: John Benjamins.
- Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45(1), 81–105.
- Segalowitz, N., & Freed, B. (2004). Context, contact and cognition in oral fluency acquisition. *Studies in Second Language Acquisition*, 26, 173–199.
- Serrano, R., Llanes, A., & Tragant, E. (2011). Analyzing the effect of context of second language learning: Domestic intensive and semi-intensive courses vs. study abroad in Europe. *System*, 39, 133–143.
- Smit, U., & Dafouz, E. (Eds.). (2013). *Integrating content and language in higher education: Gaining insights into English-medium instruction at European universities*. Association Internationale de Linguistique Appliquée/International Association of Applied Linguistics (AILA) 25. Amsterdam/Philadelphia: John Benjamins.
- Sunderman, G., & Kroll, J. (2009). When study abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics*, 30, 1–21.
- Taguchi, N. (2008). Cognition, language contact, and the development of pragmatic comprehension in a study-abroad context. *Language Learning*, 58(1), 33–71.
- Towell, R., Hawkins, T., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17(1), 84–119.
- Tracy-Ventura, M., Dewaele, J. M., Koylu, Z., & McMannus K. (2016). Personality changes after a year abroad: A mixed methods study. *Study Abroad Research in Second Language Acquisition and International Education*, 1(1), 107–126.
- Trooboff, M., & Rayman, J. (2008). Employers attitudes toward study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 15, 17–33.
- Wächter, B., & Maiwörm, F. (2014). *Rethinking the role of English-taught programmes in the EHEA*. Brussels: ACA.
- Wilkinson, S. (1998). On the nature of immersion during study abroad: Some participant perspectives. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 4, 121–138.

Computer-Assisted SLA

Hayo Reinders and Glenn Stockwell

Background

Despite the ubiquity of technology in language learning and teaching, and a widespread interest in its potential to enhance, and potentially transform, language education, research in the area of technology-assisted second language acquisition (SLA) is both recent and relatively limited. In this chapter we first review how the field has developed, moving away from its earlier focus on demonstrating the ‘advantages’ of technology, to our current understanding of its affordances and constraints. Next, we review the relationship between SLA and computer-assisted language learning (CALL) and show how CALL research has increasingly drawn on research in SLA and, in recent years, is starting to exert its own influence on our understanding of SLA processes. In the following section we draw on the 10 principles of SLA identified by Ellis (2008) to illustrate this relationship, and conclude with a number of future directions for the field.

The use of technology in teaching languages is far from new, and language teachers have long sought to discover how emerging technologies could be effectively used to facilitate the language learning process. Early, bulky stand-alone tools in the 1980s gave way to the use of networked machines in the 1990s, which were replaced with more and more sophisticated and portable tools that allowed increased interactivity and multimedia capabilities through the 2000s and up to the present day. Modern technologies have an almost constant, stable, and fast connection to the internet in most regions, and devices such as laptop computers, smartphones, tablets, and wearable technologies have become much more affordable. These technologies bring with them different affordances, that is, different possibilities and potentialities, which means that research needs to be carried out in a range of environments to investigate the various ways in which the technology may be used for second language (L2) learning.

It is not only the technologies that have developed over time. The methods of researching these technologies have also evolved, moving from the effectiveness studies that predominated in early years through to more sophisticated studies aimed at identifying how the specific affordances of these technologies can affect the language learning process (Reinders & White, 2010). On face value, effectiveness studies do seem to have an important place in determining how technology can be used in SLA. There is a danger, however, that we fall

into the trap of the ‘burden of proof’ as cited by Burston (2003), where we feel the need to prove that using technology is more effective than not using it, due to the fact that we have invested so much time and money in its implementation. One problem with the desire to demonstrate the superiority of technology is that it has resulted in a body of research that overclaims the effectiveness of technology in SLA, in many cases with unsuitable or inadequate research designs (see Felix, 2008, for a discussion). The question of whether technology is effective in SLA still persists, and those who are new to the field will often doubt the effectiveness of technology use, a fact that has no doubt been the impetus for studies such as Grgurović, Chapelle, and Shelley’s meta-analysis (2013), which suggested that there is a small but significant effect of using technology on L2 proficiency in classroom instruction.

Given the efforts invested by those who implement technology in language learning and teaching environments, the fact that technology can have a positive impact on SLA is certainly reassuring. As yet, however, little is known about the mechanisms behind the benefits attributed to technology in this process. While *general* learning theories have always occupied a role in CALL research, the field has relied heavily on *SLA* theories (Hubbard, 2008), and as such it is not surprising that shifts in theories in SLA are often reflected in CALL as well (Levy & Stockwell, 2006). In addition, it is becoming increasingly evident that technology changes the language learning environment sufficiently that the role of technology itself must be considered in the theories that are applied in CALL (Stockwell, 2014). Theories that relate more specifically to technology use have started to be applied to CALL recently, such as situated learning (Brown, Collins, & Duguid, 1989), which focuses on the ability of mobile devices to interact with the environment, and dual coding theory (Paivio, 2007), which considers the provision of input for learners through both visual and audio codes, thereby allowing input to be processed through different channels.

Over the past several years, however, there has been an indication that studies on the role of technology can inform SLA theory as well. As an example, the use of technology problematises the distinction between learning and teaching and the notion of ‘instruction.’ Most people would probably consider the use of a news website by a classroom teacher to be a form of instruction. If that same website was used by a student not enrolled in any classes, it would probably not be considered a form of instruction. But how about a website that offers self-study language learning resources? Clearly some of the ‘instruction’ in such cases could be programmed into the website and it could be argued that a form of instruction does indeed take place. More questionable would be the case of a website designed to pair learners for a language exchange. In this case the site creates certain conditions for learning to take place but there is no actual instruction. Clearly, when it comes to technology, the lines between what does and does not constitute instruction are not clear (see Loewen, 2015). For the purposes of this chapter, however, we focus primarily on cases where technology is used for direct instruction. We include all uses of technology, including those not drawing on computers, in applying the commonly used term ‘Computer-Assisted Second Language Acquisition,’ or CASLA. In the next section, we will focus on some of the current issues that occupy the field.

Current Issues

The earlier focus on demonstrating the superiority of CALL compared with ‘traditional’ instruction has given way to an understanding (in accordance with Krashen’s first law of technology; 1986) that technology is neither beneficial nor detrimental in and of itself. Instead, the field has more recently concerned itself more with identifying when and how technology can be used to enhance learning and teaching. Reinders and White (2010) synthesise these

Table 20.1 Affordances of CALL

Organisational affordances	Improved access
	Storage and retrieval of learning behavior records and outcomes
	Sharing and recycling of materials
	Cost efficiency
Pedagogical affordances	Improved authenticity of L2 input
	Improved interaction between learners, between learners and native speakers, as well as between learners and instructor
	Situated learning (e.g., the availability of technology outside the classroom to support language use)
	The use of multimedia
	New forms of learning and teaching activities
	Nonlinearity (e.g., through hyperlinking of texts)
	Alternative forms of (giving and receiving) feedback
	Monitoring and recording of learning behavior and progress
	Greater control over the learning process
	Empowerment of learners and teachers by enabling them to make independent choices about their own learning

‘affordances,’ or potential, contextually determined, and contextually dependent benefits of using technology, and distinguish between organisational and pedagogical affordances. The results of their study are summarised in Table 20.1.

The organisational affordances relate to potential benefits for the instructional context, such as by reducing the cost of delivery (for example, when students engage in computer-supported self-study), or by making learning and teaching opportunities more widely available (for example through the use of online resources that can be accessed without time and space constraints). Pedagogical affordances include the ability to provide opportunities for situated learning (i.e., learning in context, for example through the use of mobile devices), opportunities for supporting learning in ways not previously possible (such as through online monitoring of student progress) and by enabling learners to control different aspects of the learning process directly (for example by determining the sequence, pace, and method of learning). However, Reinders and White (2016) argue that the realisation of such affordances depends on local factors; for example in the case of learner control and empowerment, technology has in many cases not had a significant impact because its transformative potential has not been realised due to other aspects of the learning and teaching ecology not allowing a significant shift in learners’ and (mostly) teachers’ expectations about the role of formal education. In other words, understanding the impact that technology can have on language acquisition depends on a deep understanding of all factors involved. This is the focus of the next section of this chapter.

Empirical Evidence

The large body of research built up in the field of CALL over the past several decades is testimony to the interest in the use of technology in the process of acquiring different aspects of a L2, including reading (Chun, 2006), writing (Kessler, Bikowski, & Boggs, 2012), listening

(Jones, 2003), speaking (Valle, 2005), vocabulary (Fuente, 2003), grammar (Sauro, 2009) and so forth. Overviews of research in this area may be found in Levy and Stockwell (2006), Stockwell (2012) and Thomas, Reinders, and Warschauer (2013), and reveal the sophistication of the range of studies carried out. The research varies widely not only in the technologies and the underlying pedagogies used, but also in the focus of the research itself, including attitudes to technology (e.g., Ayres, 2002; Winke & Goertler, 2008), patterns of engagement (e.g., Milligan, Littlejohn, & Margaryan, 2014), and, of course, acquisition of different aspects of a L2.

The results of these studies have also been quite varied, an outcome that is hardly surprising considering the complexities and variables involved in the learning process. Furthermore, empirical measures of SLA in both CALL and non-CALL contexts are typically limited to one or two specific skills or areas that can be measured through the instruments that are used, meaning that individual studies can give us only a glimpse into certain smaller aspects of the larger phenomenon of L2 learning. It is also important to note that, as pointed out by Felix (2005), focusing only on the outcomes of research into SLA through CALL is unlikely to give a clear picture of how and why learning takes place, and there is a need to also investigate the processes of learning in order to understand more fully the role that technology may play, hence the recent interest in research syntheses and meta-analyses in this area (e.g., Grgurović et al., 2013; Sauro, 2011).

An important area of research is the provision of (conditions for) interaction with other people through various forms of computer-mediated communication (CMC). Research into CMC for language learning has undergone transformations that largely follow technical developments, and have included text chat (Lai & Zhao, 2006), email (Stockwell & Harrington, 2003), audioconferencing and videoconferencing (Wang, 2004), and more recently, social networking (Mok, 2012). Other forums that have allowed interaction between students and their interlocutors have included virtual worlds (Toyoda & Harrison, 2002) and video games (Peterson, 2012). Each of these forms of CMC brings different combinations of the affordances listed in the previous section, and has the potential to impact different language skills and areas as a result of the mode of communication (i.e., textual, visual, and/or oral), and the degree of synchronicity (e.g., synchronous videoconferencing vs. asynchronous email). Studies have shown that communication through CMC bears a number of similarities to face-to-face language, specifically in terms of the presence of negotiation of meaning. As Bower and Kawaguchi (2011) point out, however, the textual nature of many forms of CMC tends to make learners more likely to notice differences between the language that they produce and that of their interlocutors, and this may enhance opportunities for acquiring the target language.

Non-CMC language learning activities have typically seen the role of the computer as either a tutor or a tool (Levy, 1997). In a tutor role, the technology provides feedback to learners based on their output, and there is a teaching presence based on some form of instructional design that is evident in the way that material is presented to the learner and in the nature of the feedback provided. Studies of this type have included investigations of simple online authoring activities such as Hot Potatoes (Shawback & Terhune, 2002) at one end of the spectrum, through to Intelligent CALL systems that analyse and adapt to individual learner abilities (Heift, 2013) at the other end. There have been a number of studies that have shown positive outcomes from learners engaging in CALL-based activities; and although there has tended to be a stronger focus on areas such as vocabulary, speech recognition software, grammar, listening, and reading, recent years have seen a steady increase in work in other more production-based areas such as writing (Chen & Cheng, 2008) and

speaking (Elimat & Abu Seileek, 2014) as well. As mentioned earlier, research has moved away from simply determining whether or not CALL is as effective or more effective than non-CALL; instead more recent research has been concerned with identifying the individual attributes of CALL that are more likely to lead to SLA. Studies such as the foregoing have suggested that learners will benefit from having sufficient feedback that can help them to target problem areas, and that having options to suit different learner styles means that these tools can be more useful to a wider range of learner proficiencies, language learning styles, and learner goals.

Thus, technology has been used in an enormous range of ways to take on a mediating role between interlocutors, a teaching role where it evaluates learner output and provides feedback, and a utilitarian role serving as to support the learning process. The effectiveness of technology in promoting L2 acquisition depends on a number of interrelated factors, but it is possible to consider several principles that are likely to lead to enhanced opportunities for learners, as described in the following section.

Pedagogical Implications

For this section we draw on the list of principles by Ellis (2005, 2008) in which he proposes 10 ‘generalisations’ of research findings from SLA studies that language educators can use as the basis for classroom instruction. We use these principles as a starting point to review studies in CASLA that have been carried out in these areas.

Principle 1: Instruction needs to ensure that learners develop both a rich repertoire of formulaic expressions and a rule-based competence.

One of the closest links between technology and SLA research has been through the development and analysis of corpora. The use of fast computers has enabled the identification of chunks or formulaic expressions that occur frequently in native-speaker language, and this has informed both the development of instructional materials and the types of language that classroom teachers introduce and assess (Granger, Gilquin, & Meunier, 2015). Learner corpora have given insight into the way that learner differences impact acquisition, and also how language develops over time (Myles, 2007). In addition, learner-generated corpora can raise student awareness and independence. By guiding learners to search, analyse and/or create corpora, common patterns of language use can be identified, as well as their underlying rules discovered.

Principle 2: Instruction needs to ensure that learners focus predominantly on meaning.

Perhaps the most widely acknowledged contribution of CASLA research has been in the area of CMC where chat transcripts and other forms of online communication (e.g., videoconferencing and the use of virtual worlds) have been extensively investigated, drawing on theories of SLA. More recently, researchers have also started to explore communication in social networks (Tran, 2016) and digital games (Cornillie, Clarebout, & Desmet, 2012). Findings confirm the importance of a focus on meaning on SLA and the ways in which the affordances of different forms of online communication (e.g., synchronous vs. asynchronous, written vs. spoken), different task conditions (with or without time pressure, with or without access to resources such as online dictionaries, etc.), affect learning outcomes (Lamy & Hampel, 2007; Sauro, 2011). An increasingly large body of

research now also exists that shows the role of technology in facilitating meaningful and meaning-focused interaction outside the classroom (see Benson & Reinders, 2011 for a compendium of such research).

Principle 3: Instruction needs to ensure that learners also focus on form.

The ability for technology to allow focus on form has long been cited as a potential benefit for language learning (Warschauer, 1996), and it is not surprising that there has been a good deal of research investigating the modes and nature of feedback that enable learners to focus on form. A recent in-depth discussion of the issue of feedback and focus on form has been carried out by Ware and Kessler (2013), who outline three modes through which feedback can be provided to learners. The first of these is face-to-face, where feedback is provided by either the teacher or peers directly to the learner based on their digital output, such as writing in a word processor or participation in chat. In other words, although the output is created digitally, the feedback from the teacher or peers on this digital output is provided to the learner face-to-face. The second mode is through human feedback that is delivered electronically. As with the previous mode, this feedback is provided by either the teacher or peers, but this time the feedback is provided through means such as chat, email or a learning management system, rather than directly face-to-face. While learners engaged in communication with others through CMC typically focus on meaning rather than form (Bower & Kawaguchi, 2011), the shift can be moved somewhat more toward form in tandem learning (e.g., Kabata & Edasawa, 2011). The degree of synchronicity has also been shown to have an impact on the degree to which learners focus on form, with synchronous types of communication such as chat being more lexically focussed than asynchronous forms of communication such as email (Stockwell, 2010). The third type of feedback that Ware and Kessler (2013) describe is computer-generated feedback. This refers to feedback that can provide automated scoring for quiz-type activities for vocabulary (e.g., Stockwell, 2007) or grammar (Heift, 2003), evaluation of writing (Chen & Cheng, 2008), speech recognition software (Elimat & Abu Seileek, 2014) or automatic transcription software (Bonneau & Colotte, 2011) that can be used in pronunciation training. Thus, technology enables focusing on form to be achieved through activities targeting specific areas of the L2 such as syntax, lexicon or pronunciation that are automatically scored and evaluated, as well as through direct teacher intervention during CALL-based tasks and activities or through computer-mediated interaction with the teacher or other learners.

Principle 4: Instruction needs to focus on developing implicit knowledge of the second language while not neglecting explicit knowledge.

Ellis and Shintani (2014) conclude that “instruction needs to be directed at developing *both* implicit and explicit knowledge, giving priority to the former” (p. 23). In other words, there is a need to provide opportunities for learners to develop their knowledge of vocabulary and grammar, while at the same time having sufficient opportunities for natural interactions, which has been argued may play a role in developing implicit knowledge (e.g., DeKeyser, 2003). There have been a few recent attempts at examining how technology can be used in developing both implicit and explicit knowledge. AbuSeileek and Abualsha’r (2014), for example, looked at how different types of computer-generated feedback could promote learners’ language development through writing essays, while Andringa and Curcic (2015) examine the role of explicit instruction on how learners process L2 information

online. They provided a brief explanation of a grammatical rule to approximately half the subjects in their study, and found a positive impact of this explicit instruction on syntactical acquisition.

Principle 5: Instruction needs to take into account the learner's built-in syllabus.

Ellis (2005) suggests that learners are more likely to acquire a L2 more effectively and efficiently if they receive instruction that is “compatible with the natural processes of acquisition” (p. 15). In order to determine individual learners’ developmental levels, teachers typically need to make assumptions about where learners might be in their built-in syllabus, or alternatively teachers need to provide broad enough language input that learners can extract the input that suits their needs. Technology has the potential to provide opportunities for learners at different points in their development through the provision of multiple pathways (Ros i Solé & Mardomingo, 2004). Using technology can allow learners to undertake activities in a nonlinear fashion, where the content can be covered in an order that suits the learners’ own individual needs and preferences. Therefore, learners can make choices in the learning process in a way that gives them freedom to undertake learning depending on their own built-in syllabus. While of course learners may not be explicitly aware of where they are in their own development, they will likely have a sense of what they feel is too difficult or too easy, and as such may be able to decide on engaging in content that they perceive as being appropriate to their learning needs. The way in which these choices are made available to learners is, of course, very dependent upon the instructional design, and there is a need to bear in mind this important affordance when designing applications for CALL, and capitalise upon it as much as possible.

Principle 6: Successful instructed language learning requires extensive second language input.

The advent of the internet in the early 1990s had an enormous impact on the availability of authentic input in the L2. Known retrospectively as Web 1.0, this resource typically took the form of static web pages in the initial stages, such as news and other informational sites, making it possible for learners to have access to large quantities of authentic target language input. Of course, one limitation with this type of authentic input is that it is targeted at native speakers, and as such it is often too difficult for learners of a lower or intermediate level of proficiency. Nonetheless, there are resources available in many languages (albeit predominantly English) that have been simplified for language learners. One example of this is BBC Learning English, which provides a simplified version of news and human interest reports, along with learner support, such as vocabulary glosses and subtitles, in various other languages. A major goal that remains for teachers is designing learning activities that take advantage of the large range of authentic materials that are available in order to have sufficient input that is appropriate to the learners’ proficiency levels.

Technology affords other sources of L2 input as well. The most obvious of these is CMC, where learners can receive input that is delivered through multiple modes and that is modified during interaction, depending on communication needs (Blake, 2000). This language input can be either textual (i.e., text chat or email) or aural (i.e., audio- or videoconferencing), and as such learners can develop both reading and listening skills depending on the type of CMC they are engaged in. A benefit that has been cited for textual forms of CMC is that they can allow the learner to focus more on language input and output in that they have time

Hayo Reinders and Glenn Stockwell

to use tools such as dictionaries in processing the content of a received message, and at the same time can review the content of a message before sending it to the interlocutor (Blake, 2013). Video and audio conference, in contrast, place a greater burden on learners to process input and to produce output in real time, and thus have been shown to result in a greater lexical focus and are more suited to learners of a higher proficiency (Stockwell, 2010). Added to this is the dimension of multimodality, where textual, oral, and even graphic modes of CMC may be used in a single communication act (Hampel & Hauck, 2006). The use of multiple modes makes it possible for learners to activate different knowledge bases that can assist in facilitating acquisition of the L2.

Principle 7: Successful instructed language learning also requires opportunities for output.

The role of the internet has changed significantly in recent years, largely as a result of the emergence of Web 2.0, which enables individuals to not only access information from the internet, but also to post information and to communicate with one another using various CMC tools. One of the primary advantages of these recent developments is that it makes it far easier for individuals to make contact with others, regardless of geographical location. Communication can take place on a one-to-one basis, such as through email, messaging or video chat, but technology can also enable information to be disseminated to a larger, and at times unknown, audience as well. As described earlier, CMC has been widely cited in CASLA research, and there are various tools that can be used to provide learners with opportunities to produce both written and oral output. There have been demonstrated differences in the quantity and quality of the language produced through synchronous CMC and asynchronous CMC, where synchronous CMC tends to be syntactically simpler and more lexically focused than asynchronous CMC (Stockwell, 2010). How to capitalise pedagogically on these differences, however, remains a challenge for teachers.

These developments have also made it possible to post information that can be accessed by a larger audience, through such forums as blogs (e.g., Pinkman, 2005) and wikis (e.g., Kessler & Bikowski, 2010), and research into blogs and wikis has indicated that learners have experienced motivational advantages through communicating to an authentic audience. The last few years have also seen an increase in the use of social networking as a potential forum for facilitating learner output as well, and more studies are appearing that examine not only the nature of learner output in these forums, but have also made it apparent that there are cultural factors that need to be kept in mind regarding the acceptance of technology in different cultural environments (e.g., Mok, 2012; Liu, 2013). Needless to say, however, technology has opened up the classroom to allow communication to extend beyond just between fellow students and the teacher to a range of interlocutors, providing opportunities for both oral and written language output in varied genres and contexts.

Principle 8: The opportunity to interact in the second language is central to developing second language proficiency.

The importance of interaction for SLA is widely recognised (Gass & Mackey, 2015) and numerous studies have demonstrated the benefits of negotiation, the provision of negative feedback, and the meaning-oriented nature of L2 interaction, among others.

Many technology-mediated environments are predicated on the notion of social interaction, with social networks being the most visible example. Participation in social networks has been shown to increase students' sense of ownership, meaningful interaction, and identity-building, as well as students' motivation (Mills, 2011; Toetenel, 2014), as has the impact of the interaction in digital games (Peterson, 2012; Reinders & Wattana, 2015). It appears digital games increase students' willingness to communicate, and the amount and range of language they produce as a result. Another, much longer established form of interaction is afforded through online language exchanges, whereby CMC tools enable L2 learners to connect with other L2 learners, or—more commonly—where L2 learners can connect with native speakers of another language, who in turn are learning their interlocutors' first language. This type of interaction has been shown to have benefits for both language acquisition, as well as the development of intercultural competence (Lamy & Hampel, 2007).

Principle 9: Instruction needs to take account of individual differences in learners.

CALL has long been used to personalise instruction to learners, in order to take individual differences into account. Where classroom instruction is necessarily limited in the ways it can cater to learners with different backgrounds, aptitudes, interests, and so on, CASLA resources can be used to (1) identify such differences and (2) tailor instruction accordingly. While early predictions, particularly in the area of iCALL, or *intelligent CALL*, claiming that computers (at that time) would take over most language instruction, have been proven to be overblown, some definite advances have been made.

In particular in the area of language testing, computer-adaptive testing, where learners' responses to previous items determine the difficulty of subsequent ones, has now come to be used widely in language testing (Tseng, 2016). Similarly, computerised diagnostic tests (which may or may not use adaptivity) are able to quickly determine a learner's approximate level (Poehner, Zhang, & Lu, 2015).

In terms of social and affective differences impacting on learning, CASLA has been used to support learners in manipulating their learning experiences based on their own preferences, and to guide them in developing the skills necessary to do so, thus providing both the 'learner training' and 'flexibility' Ellis and Shintani (2014) refer to. Online self-access resources (Reinders & Darasawang, 2012) allow learners to take some degree of control over their learning while still being guided. A similar approach is the use of Personal Learning Environments (or PLEs; Plastina, 2015), which use commonly available communication tools to support learners in goal setting, monitoring their progress, and building portfolios. An important feature of such environments is their social aspect, which allows learners to connect with peers, outside of the language classroom.

The use of CASLA resources has enabled language instruction to adopt a flipped approach (Hung, 2015) whereby classroom time is used to provide individual support, while learners work on tasks appropriate to them and prepare for classroom time either on their own or with peers. Materials and resources that can be accessed outside of the classroom and that provide automated feedback free up the teacher to work on other things. However, it could be argued that the most important contribution of CASLA to better accommodating learner differences has been to provide educators and learners with the tools to allow them to extend formal education to nonformal (related to formal education but separate from it) and informal (unrelated to formal education) spaces¹ (Benson, 2011). Through this extension, learners have access to a much wider range of learning

Hayo Reinders and Glenn Stockwell

opportunities, provided not just in the way the teacher deems appropriate, but that can be adjusted by learners themselves.

Principle 10: In assessing learners' second language proficiency, it is important to examine free as well as controlled production.

CALL can involve written or spoken language conducted either with other people (i.e., CMC) or directly with the computer. As described earlier, CMC may include text chat, email, audio chat, video-conferencing, social networking, and digital games, and the nature of the language produced will depend very much on the assigned tasks. Communication tasks through CMC would generally be considered as a forum for free language production, and there has been quite extensive research into these types of activities and their impact on SLA (e.g., Monteiro, 2014; Stockwell & Harrington, 2003; Tare et al., 2014). These studies have shown that learners engage in similar behaviours that are exhibited in face-to-face contexts, but that the mode of communication has an impact on the complexity and accuracy of the language produced.

Controlled production tends to occur when learners interact directly with the computer itself. Both written and oral production can fit into the category of pattern matching, where only a limited number of responses to a prompt are considered acceptable. These responses typically take the form of a short answer to a question, or sentence-level translation (e.g., Heift, 2003). Alternatively, interacting with the computer can also include continuous production, where language can be analysed for features such as grammaticality and style (see Ware & Warschauer, 2006). Oral production depends heavily on automatic speech recognition (ASR), which converts oral output into textual form so that it can be parsed for use in either pattern matching or continuous forms of analysis. Speech recognition is an area that shows a good deal of promise, and while there are still limitations with the accuracy of recognition of language produced by nonnative speakers (Warren, Elgort, & Crabbe, 2009), developments are occurring rapidly to overcome these difficulties (Ross, 2015).

The ways in which technology can be used to enhance L2 acquisition have shown to be broad, but the same basic principles of best practice for instructed SLA can still be applied. This is not to say that the role of technology should be ignored, but the fact that technology will necessarily make a difference to the overall learning environment must be kept in mind (Levy, 2000; Stockwell, 2012). That is to say, that when technology is introduced into the equation, it will have some impact on the ways in which learners interact with the content, with other learners, or with the teacher. In saying this, however, the ultimate aim of learning a language remains the same, and technologies can be used to facilitate this provided instruction takes into consideration the affordances of the technology and the environment.

Future Directions

There are three broad areas where technology is likely to have a significant influence on the way people learn languages in the coming years, and where there exists an urgent need for research to understand how learners interact with and can benefit from the technologies that are being used in their language learning contexts. Rather than attempting to pinpoint the always-changing technologies, instead in this section we identify three broad areas of affordances that new developments offer.

Mobility

Probably the most developed of the three areas is our understanding of the benefits of mobility on language acquisition. Reinders and Pegrum (2016) propose a framework for evaluating MALL (mobile-assisted language learning) resources and identify a range of affordances, such as their ability to extend learning beyond the classroom, the opportunities for social interaction, and options for personalising learning, among others. What is not well understood, however, is how learners use mobile resources for the purpose of learning, and how teachers can best support learners in this endeavor.

Augmentation

Relevant in the context of MALL as well as more broadly in education in general, Atkinson (2010) cites Semin and Cacioppo (2008, p. 140) as saying that “a sea change in research and theory” has occurred where now much greater recognition exists of embodied, extended, and distributed forms of cognition. The former sees cognition as grounded in and intricately linked to bodily movements and states. Extended cognition (Clark & Chalmers, 1998) sees an interdependence between the mind and its environment. Distributed cognition further recognises that knowledge can be held in networks, with each element in a network having access to the knowledge but only in relation to other elements in that network, resulting in greater efficiency (Clark, 2008). Theories of embodied, extended, and distributed cognition offer an alternative to cognitivist views of language acquisition. As learners have ever-increasing access to tools and resources to help them acquire and use the language, this is likely to have a significant impact on how (and even *if*) languages are learned (in particular as machine translation, natural language processing, and related technologies become more powerful). Mobile technologies, for example, with their affordance for situated learning, allow learners to be offered context-specific vocabulary, or pragmatically appropriate conversational language (Pegrum, 2014). The use of touch and gestures for interacting in CALL can also be beneficial for language learning (Reinders, 2014), and haptic feedback has potential for providing alternative forms of input enhancement and correction (Reinders, 2014). Virtual and augmented reality tools enable the seamless combination of physical and digital resources, so that, for example objects in a room can be ‘annotated’ with their foreign language translation, as learners interact with them, wearing headsets or other forms of wearable computing.

Ubiquity

There is considerable discussion at present about the potential for disruption from ‘the internet of things’ (IOT) and related technologies. IOT refers to the connection of physical devices, such as cars, fridges, syringes, and door handles, to the internet, and estimates range from 20–100 billion connected devices by 2020 (Evans, 2011). The first applications are starting to be seen in health, for example by monitoring outpatients’ medicine intake or tracking the location of equipment in hospitals. The potential of IOT for education is only just starting to be explored with the first projects looking at the ways in which rooms can recognise learners and track attendance, and where items such as books can record and report usage and achievement, or even adjust content depending on performance or the location where the learner is at a given time.

What all these areas have in common is that they extend language learning beyond the classroom, as well as beyond formal education. As a result, it is likely that learning

will become not only more of a lifelong (spanning one's lifetime) but also a lifewide (not confined to a particular location, such as a school) activity. Technology will increasingly allow learners to gain access to learning opportunities that are not only increasingly varied, but also increasingly connected to other learners, and increasingly individualised. The impact of these developments on SLA offers a fascinating and as yet underexplored field of research.

Note

1. Benson distinguishes between these as follows: "non-formal education often refers to classroom or school-based programmes that are taken for interest and do not involve tests or qualifications, while informal education refers more to non-institutional programmes or individual learning projects" (2011, p. 10).

References

- AbuSeileek, A., & Abualsha'r, A. (2014). Using peer computer-mediated corrective feedback to support EFL learners' writing. *Language Learning & Technology*, 18(1), 76–95.
- Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing. *Studies in Second Language Acquisition*, 3 (2), 237–268.
- Atkinson, D. (2010). Extended, embodied cognition and second language acquisition. *Applied Linguistics*, 31(5), 99–622.
- Ayres, R. (2002). Learner attitudes towards the use of CALL. *Computer Assisted Language Learning*, 15(3), 241–249.
- Benson, P. (2011). *Teaching and researching autonomy in language learning*. Harlow: Pearson Longman.
- Benson, P., & Reinders, H. (Eds.). (2011). *Beyond the language classroom*. Basingstoke: Palgrave Macmillan.
- Blake, R. (2000). Computer mediated communication: A window on L2 Spanish interlanguage. *Language Learning & Technology*, 4(1), 120–136.
- Blake, R. (2013). *Brave new digital classroom: Technology and foreign language learning* (2nd ed). Washington, DC: Georgetown University Press.
- Bonneau, A., & Colotte, V. (2011). Automatic feedback for L2 prosody learning. In I. Ipsic (Ed.), *Speech and language technologies* (pp. 55–70). Rijeka, Croatia: Intech.
- Bower, J., & Kawaguchi, S. (2011). Negotiation of meaning and corrective feedback in Japanese/English eTandem. *Language Learning & Technology*, 15(1), 41–71.
- Brown, J. S., Collins, A., & Duguid, S. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Burston, J. (2003). Proving IT works. *CALICO Journal*, 20(2), 219–226.
- Chen, C. F. E., & Cheng, W. Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112.
- Chun, D. M. (2006). CALL technologies for L2 reading. In L. Ducate & N. Arnold (Eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching* (pp. 81–98). San Marcos, TX: CALICO.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York: Oxford University Press.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58, 10–23.
- Cornillie, F., Clarebout, G., & Desmet, P. (2012). Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. *ReCALL*, 24(3), 257–278.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313–349). Oxford: Blackwell.

- Elimat, A. K., & Abu Seileek, A. F. (2014). Automatic speech recognition technology as an effective means for teaching pronunciation. *The JALT CALL Journal*, 10(1), 21–47.
- Ellis, R. (2005). Principles of instructed language acquisition. *Asian EFL Journal*, 7(3), 9–24.
- Ellis, R. (2008). Principles of instructed second language acquisition. *CAL Digest*, 1–6. Available from <http://www.cal.org/resource-center/briefs-digests/digests>
- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Abingdon: Routledge.
- Evans, D. (2011). The internet of things: How the next evolution of the internet is changing everything: Cisco white paper. Retrieved from http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf
- Felix, U. (2005). Analysing recent CALL effectiveness research—towards a common agenda. *Computer Assisted Language Learning*, 18(1–2), 1–32.
- Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learned in two decades of research? *ReCALL*, 20(2), 141–161.
- Fuente, M. J. (2003). Is SLA interactionist theory relevant to CALL? A study on the effects of computer-mediated interaction in L2 vocabulary acquisition. *Computer Assisted Language Learning*, 16(1), 47–81.
- Gass, S., & Mackey, A. (2015). Input, interaction and output in second language acquisition. In J. Williams & B. VanPatten (Eds.), *Theories in second language acquisition* (2nd ed., pp. 180–206). Mahwah, NJ: Erlbaum.
- Granger, S., Gilquin, G., & Meunier, F. (Eds.). (2015). *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press.
- Grgurović, M., Chapelle, C. A., & Shelley, M. C. (2013). A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL*, 25(2), 165–198.
- Hampel, R., & Hauck, M. (2006). Computer-mediated language learning: Making meaning in multi-modal virtual learning spaces. *The JALT CALL Journal*, 2(2), 3–18.
- Heift, T. (2003). Multiple learner errors and meaningful feedback: A challenge for ICALL systems. *CALICO Journal*, 20(3), 533–548.
- Heift, T. (2013). Clicking for help. In P. Hubbard, M. Schultz, & B. Smith (Eds.), *Human-computer interaction in language learning: Studies in honor of Robert Fischer* (CALICO Monograph Series, pp. 187–202). San Marcos, TX: CALICO.
- Hubbard, P. (2008). Twenty-five years of theory in the CALICO Journal. *CALICO Journal*, 25(3), 387–399.
- Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning*, 28(1), 81–96.
- Jones, L. C. (2003). Supporting listening comprehension and vocabulary acquisition with multimedia annotations: The students' voice. *CALICO Journal*, 21(1), 41–65.
- Kabata, K., & Edasawa, Y. (2011). Tandem language learning through a cross-cultural keypal project. *Language Learning & Technology*, 15(1), 104–121.
- Kessler, G., & Bikowski, D. (2010). Developing collaborative autonomous learning abilities in computer mediated language learning: Attention to meaning among students in wiki space. *Computer Assisted Language Learning*, 23(1), 41–58.
- Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. *Language Learning & Technology*, 16(1), 91–109.
- Kranzberg, M. (1986). Technology and history: Kranzberg's laws. *Technology and Culture*, 27(3), 544–560.
- Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology*, 10(3), 102–120.
- Lamy, M. N., & Hampel, R. (2007). *Online communication in language learning and teaching*. Basingstoke: Palgrave Macmillan.
- Levy, M. (1997). *Computer assisted language learning: Context and conceptualization*. Oxford: Pergamon Press.

- Levy, M. (2000). Scope, goals and methods in CALL research: Questions of coherence and autonomy. *ReCALL*, 12(2), 170–195.
- Levy, M., & Stockwell, G. (2006). *CALL Dimensions: Options and issues in computer assisted language learning*. Mahwah, NJ: Lawrence Erlbaum.
- Liu, J.Y. (2013, May). *Using Facebook for learning writing: EFL students' perceptions and challenges*. Paper presented at the 30th International Conference on English Teaching and Learning, Tainan, Taiwan.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Milligan, C., Littlejohn, A., & Margaryan, A. (2014). Patterns of engagement in connectivist MOOCs. *MERLOT Journal of Online Learning and Teaching*, 9(2), 149–159.
- Mills, N. (2011). Situated learning through social networking communities: The development of joint enterprise, mutual engagement, and a shared repertoire. *CALICO Journal*, 28(2), 45–368.
- Mok, J.C.H. (2012). Facebook and learning: Students' perspectives on a course. *Journal of the NUS Teaching Academy*, 2(3), 131–143.
- Monteiro, K. (2014). An experimental study of corrective feedback during video-conferencing. *Language Learning & Technology*, 18(3), 56–79.
- Myles, F. (2007). Investigating learner language development with electronic longitudinal corpora: Theoretical and methodological issues. In L. Ortega & H. Byrnes (Eds.), *The longitudinal study of advanced L2 capacities* (pp. 58–72). Hillsdale, NJ: Lawrence Erlbaum.
- Paivio, A. (2007). *Mind and its evolution: A dual coding theoretical approach*. Mahwah, NJ: Erlbaum.
- Pegrum, M. (2014). *Mobile learning: Languages, literacies and cultures*. Basingstoke: Palgrave Macmillan.
- Peterson, M. (2012) Language learner interaction in a massively multiplayer online role-playing games. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 70–92). Basingstoke: Palgrave Macmillan.
- Pinkman, K. (2005). Using blogs in the foreign language classroom: Encouraging learner independence. *The JALT CALL Journal*, 1(1), 12–24.
- Plastina, A.F. (2015). Do-it-our-way or do-it-yourself? ESP learner control in personal learning environments. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 5(1), 35–52.
- Poehner, M.E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337–357.
- Reinders, H. (2014). Touch and gesture-based language learning: Some possible avenues for research and classroom practice. *Teaching English with Technology Journal*, 14(1), 3–8.
- Reinders, H., & Darasawang, P. (2012). Diversity in language support. In G. Stockwell (Ed.), *Computer-assisted language learning: Diversity in research and practice* (pp. 49–70). Cambridge: Cambridge University Press.
- Reinders, H., & Pegrum, M. (2016). Supporting language learning on the move. An evaluative framework for mobile language learning resources. In B. Tomlinson (Ed.), *Second language acquisition research and materials development for language learning* (pp. 221–233). London: Taylor & Francis.
- Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, 27(1), 38–57.
- Reinders, H., & White, C. (2010). The theory and practice of technology in materials development and task design. In N. Harwood (Ed.), *Materials in ELT: Theory and practice* (pp. 58–80). Cambridge: Cambridge University Press.
- Reinders, H., & White, C. (2016). Twenty years of learner autonomy and technology: How far have we come and where to next? *Language Learning & Technology*, 20(2), 143–154.
- Ros i Solé, C., & Mardomingo, R. (2004). Trayectorias: A new model for online task-based learning. *ReCALL*, 16(1), 145–157.

- Ross, G. (2015, October). *Talking to your computer: Using web speech to develop student communication skills and autonomy*. Paper presented at the KOTESOL 2015 International Conference, Seoul.
- Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology*, 13(1), 96–120.
- Sauro, S. (2011). SCMC for SLA: A research synthesis. *CALICO Journal*, 28(2), 369–391.
- Semin, G. R., & Cacioppo, J. T. (2008). Grounding social cognition: Synchronization, coordination, and co-regulation. In G. R. Semin & E. R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 119–147). New York: Cambridge University Press.
- Shawback, M., & Terhune, M. (2002). Online interactive courseware: using movies to promote cultural understanding in a CALL environment. *ReCALL*, 14(1), 85–95.
- Stockwell, G. (2007). Vocabulary on the move: Investigating an intelligent mobile phone-based vocabulary tutor. *Computer Assisted Language Learning*, 20(4), 365–383.
- Stockwell, G. (2010). Effects of multimodality in computer-mediated communication tasks. In M. Thomas & H. Reinders (Eds.), *Task-based language teaching and technology* (pp. 83–104). London & New York: Continuum Books.
- Stockwell, G. (2012). Introduction. In G. Stockwell (Ed.), *Computer-assisted language learning: Diversity in research and practice* (pp. 49–70). Cambridge: Cambridge University Press.
- Stockwell, G. (2014). Exploring theory in computer-assisted language learning. In X. Deng & R. Seow (Eds.), *Alternative pedagogies in the English language & communication classroom: Selected papers from the fourth CELC Symposium for English Language Teachers* (pp. 25–30). Singapore: Centre for English Language Communication, National University of Singapore.
- Stockwell, G., & Harrington, M. (2003). The incidental development of L2 proficiency in NS-NNS email interactions. *CALICO Journal*, 20(2), 337–359.
- Tare, M., Golonka, E. M., Vatz, K., Bonilla, C. L., Crooks, C., & Strong, R. (2014). Effects of interactive chat versus independent writing on L2 learning. *Language Learning & Technology*, 18(3), 208–227.
- Thomas, M., Reinders, H., & Warschauer, M. (Eds.). (2013). *Contemporary computer-assisted language learning*. New York: Continuum.
- Toetenel, L. (2014). Social networking: A collaborative open educational resource. *Computer Assisted Language Learning*, 27(2), 149–162.
- Toyoda, E., & Harrison, R. (2002). Categorization of text chat communication between learners and native speakers of Japanese. *Language Learning & Technology*, 6(1), 82–99.
- Tran, P. (2016). Training learners to use Quizlet vocabulary activities on mobile phones in Vietnam with Facebook. *The JALT CALL Journal*, 12(1), 43–56.
- Tseng, W. T. (2016). Measuring English vocabulary size via computerized adaptive testing. *Computers & Education*, 97, 69–85.
- Valle, L. M. (2005). Analyzing oral skills in voice e-mail and online interviews. *Language Learning & Technology*, 9(3), 146–163.
- Wang, Y. (2004). Supporting synchronous distance language learning with desktop videoconferencing. *Language Learning & Technology*, 8(3), 90–121.
- Ware, P., & Kessler, G. (2013). CALL and digital feedback. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary computer-assisted language learning* (pp. 323–339). London: Bloomsbury.
- Ware, P., & Warschauer, M. (2006). Electronic feedback and second language writing. In K. Hyland & F. Hyland (Eds.), *Feedback on ESL writing: Context and issues* (pp. 105–122). Cambridge: Cambridge University Press.
- Warren, P., Elgort, I., & Crabbe, D. (2009). Comprehensibility and prosody ratings for pronunciation software development. *Language Learning & Technology*, 13(3), 87–102.
- Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3–20). Tokyo: Logos International.
- Winke, P., & Goertler, S. (2008). Did we forget someone? Students' computer access and literacy for CALL. *CALICO Journal*, 25(3), 482–509.



Taylor & Francis

Taylor & Francis Group
<http://taylorandfrancis.com>

Section V

Individual Differences and Instructed Second Language Acquisition



Taylor & Francis

Taylor & Francis Group
<http://taylorandfrancis.com>

Social Dimensions and Differences in Instructed SLA

Patricia A. Duff

Background

For many individuals, language learning is a purely intellectual activity, something they are passionate about because they enjoy understanding how different languages “work.” They may be quite satisfied learning languages on their own from print or online materials with few or no expectations of ever using the language in social situations. Reading or producing texts in the target language may be sufficiently rewarding for many such learners. Hyper-polyglots often fall into this category (e.g., Erard, 2012). Others may learn languages only because it is an academic requirement and they, too, may have few opportunities or ambitions to use the language with others, either when learning it or later. Social dimensions that they, as individuals or groups, bring to the activity of learning may be of little consequence.

However, many language learners develop proficiency in another language in a much more obviously social context and with social engagement and participation in particular discourse communities as both a means and an end to their second (or additional) language acquisition (SLA). For these learners, social aspects of their lives and linguistic engagements may be quite consequential. Interactions with others can scaffold, mediate, and motivate their learning, whether with teachers and classmates in classrooms, in extracurricular discussions with roommates and host families in study abroad contexts, or in interactions in other linguistic sites in their daily lives (Duff & Surtees, in press; Swain & Deters, 2007). Conversely, exclusion from participation to their full potential or recognition as legitimate and valued class members during classroom discussions or group work based on real or perceived social aspects of their lives affects not only the quality of their language practice, but also notions of who they are—and might become—as users of the language. Their persistence with language study may also be affected. Such scenarios likely contribute to the high rate of attrition reported in many language programs after compulsory coursework has ended (e.g., Bradshaw, 2007).

Interaction in the service of language learning is both social and cognitive. Humans in interaction, particularly in face-to-face encounters, are physical beings of particular ages,¹ races, ethnicities, nationalities, genders, sexualities, religions, (im)migration/citizenship statuses, lengths of residence in a particular context, lifestyles, employment types, and

socioeconomic classes, among other visible and nonvisible dimensions. These facets of learners' lives often become salient to their interlocutors as well, who have their own social attributes and preferences (or biases). Thus, the "social" is necessarily a relational aspect of learning and, unfortunately, not all variants within a descriptive category enjoy the same social standing vis-à-vis others. These attributes may be associated with "individual differences." However, unlike differences such as *learning styles*, *aptitude*, or *willingness to communicate*, they reference larger social constructs, groups, histories, boundaries, and ideologies that are also discursively invoked and (re)produced in social settings. These biological and social categories often become very relevant and influential in learners' situated SLA experiences, not because the categories are necessarily the most important aspects of their identities from their own point of view or in their L2 settings, but because learners may be positioned in particular ways by classmates, teachers, and members of society on the basis of these factors.

Although Block (2003) and Firth and Wagner (1997) are often credited with the "social turn" in SLA, awareness of the importance of social contexts, roles, identities, and intra- and intergroup dynamics began decades earlier with work in second-language (L2) pragmatics, sociolinguistics of SLA, the sociology of language learning and loss, acculturation and accommodation theories, social distance, ethnolinguistic vitality, and research in social psychology (see, e.g., reviews in Ellis, 2008). However, social factors, and aspects of learners' social *identities*, in particular, have received renewed theoretical and empirical attention by researchers in SLA in recent years (e.g., Atkinson, 2011; Batstone, 2010; Block, 2007; Menard-Warwick, 2005; Norton, 2013). The social turn (or *return* to an examination of social factors) can be attributed, in part, to the influence of various sociological and cultural (and discursive) psychological theories in SLA, with a greater emphasis on learners' *identities*, *communities*, and *trajectories* (e.g., Duff, 2012; Duff et al., 2013; Swain & Deters, 2007).

Attention to social dimensions of SLA also reflects the complex range of language contact and learning situations associated with increased globalization and mobility, on the one hand, and greater commitment to indigeneity, on the other hand. More SLA research is now being conducted than before in multilingual postcolonial contexts (both Western and non-Western); in transnational situations with intensive voluntary and involuntary migration and mobility; in an assortment of internet-mediated social networks and virtual worlds; in Indigenous and heritage-language language communities; and so on (Duff, 2015). In these contexts, which are often directly connected with classroom learning as well, social dimensions and differences among learners and the (relative) social statuses of their languages and backgrounds may be very salient (Douglas Fir Group, 2016). These new situations are different from traditional SLA research, which focused to a large extent on (1) L2 competence/performance and learning/use among middle-class international students studying English at universities in the US and other anglophone countries, or studying foreign or modern languages in schools and universities at home (e.g., Ellis, 2008); and (2) immigrant students (Generation 1.5 and others; e.g., Kim & Duff, 2012) from various backgrounds, some of them facing major linguistic and academic challenges.

This chapter examines some of the social dimensions and differences of greatest current relevance to SLA research and considers some pedagogical implications.

Current Issues

In this section, I review key social dimensions relevant to current SLA research. The following Key Concept box also provides a brief glossary of the major constructs discussed.

Key Concepts

Social turn: A shift or expansion in focus in SLA associated with sociocultural theories and (other) interdisciplinary approaches. The goal is to better understand the social contexts and sociocognitive dimensions of learners' lives, investments, interactions, and meanings in relation to language learning.

Identity: Aspects of students' (and instructors') lives that are meaningful to them and/or others in relation to their social worlds and histories.

Community: The groups of people in which students live, study, work, or play or aspire to do so.

Trajectories: The pathways or direction individuals take (or their learning exhibits) as they move through life, typically described in terms of lines, curves, or arcs (upward or downward, in relation to their goals, or notions of progress or achievement).

Social networks, individual networks: The associations learners have with others in their immediate and distributed social circles, and the strengths of relationships or ties with each person or cluster, and the role of these relationships in supporting L2 learning and use.

Structure and agency: Two complementary and interacting dimensions of social life based on social structures and conventions into which people are socialized, on the one hand, and individuals' abilities and autonomy to take action in pursuit of their own goals, which are nevertheless mediated or constrained by such social structures, histories, and relationships.

Intersectionalities: The amplifying interactions between two or more social categories or designations, such as gender + race, or sexuality + social class, within a particular sphere of social life (such as SLA).

Social class: The socioeconomic histories and circumstances of individuals, based on such factors as their families' or their own educational backgrounds, employment, property, consumption patterns, networks, income, material possessions and resources, and other forms of capital (social, cultural, economic, symbolic) associated with different (stratified) levels of power and prestige in society.

Social Structure and Agency in SLA: Deconstructing Categories and Labels

In earlier SLA research, social categories were often considered stable, durable, shared, and operationalizable group variables (e.g., *French* vs. *English*, in Canadian research looking at attitudes and motivation toward the two languages and their speakers in English vs. French Canada; see review in Duff, 2012). Social categories people identified with, whether based on ethnicity, race, gender, or occupation, were not considered *individual differences* people brought to their learning but rather represented “social aspects of SLA” (Ellis, 2008). In contrast, constructivist and poststructural approaches to language learning and education view these categories, and related identities, as both individual and social. In addition, current theory does not view these categories as fixed (e.g., binaries or closed sets) or static, or always relevant in a learning situation. Rather, they are invoked in certain situations, taken up, performed, and resisted in dynamic ways in actual contexts of learning and using language (Block, 2003, 2007; Menard-Warwick, 2005; Norton, 2013). For example, in Abdi's (2011) research of Spanish language teaching and learning in a Canadian high school with a mixture of students from Spanish (Latina/o) and non-Spanish backgrounds, the teacher made assumptions about whether certain students came from Spanish heritage backgrounds

Patricia A. Duff

or not. She asked students to answer questions to model responses for others accordingly. However, sometimes her assumptions about students' backgrounds and their oral or written proficiency in Spanish were incorrect (e.g., when assuming a Brazilian student was from a Spanish-speaking background). Therefore, she was constructing or foregrounding certain identities or social categories for students in the class (and did so for herself in particular ways as well), not all of which were the identities students themselves chose or valued or were even accurate. Such mismatches are often discussed in SLA terms of the "social construction" of categories or identities, or resistance to them, or the uptake and performance of certain identities imposed by others or by oneself. Dealing with categories also entails "identity work," which is a fundamental aspect of learning in social contexts.

As Block (2007) writes:

Identity work occurs in the company of others—either face-to-face or in an electronically mediated mode—with whom to varying degrees individuals share beliefs, motives, values, activities and practices. Identities are about negotiating new subject positions at the crossroads of the past, present and future. Individuals are shaped by their sociohistories but they also shape their sociohistories as life goes on. The entire process is conflictive as opposed to harmonious and individuals often feel ambivalent. There are unequal power relations to deal with, around the different capitals—economic, cultural and social—that both facilitate and constrain interactions with others in the different communities of practice with which individuals engage in their lifetimes.

p. 27

Research examining *intersections* among these social categories (also known as *intersectionalities*, a term from feminist politics and sociology; e.g., McCall, 2005) underscores the potential significance not of single points or dimensions of difference (e.g., race, ethnicity, class, gender, disability) but of intersections among them—such as being poor, white, and female; or black, gay, and male—and attendant experiences of exclusion or subordination, for example. How these intersections and interactions play out in instructed SLA requires further attention (see Block & Corona, 2014; Carr & Pauwels, 2006).

Therefore, instead of binaries, continua (e.g., in terms of *masculinities*, in the plural, not singular) and intersectionalities are discussed increasingly in social science research, including SLA. "Good language learner" attributes, once considered individual, mostly (meta) cognitive features, have been critiqued and reframed because, while these attributes may be conducive to positive learning outcomes, they do not *ensure* positive learning experiences in classroom or extracurricular contexts because of potential bias, social distance, or exclusion that is often beyond learners' control, no matter how motivated and cognitively resourceful they might be (e.g., Norton, 2013; Norton & Toohey, 2001).

Another influential line of social theory relevant to SLA holds that *social structure* (e.g., stratification, power differentials, particular kinds of cultural and social capital, policies, institutions, and histories that members of society are socialized into) can both facilitate and constrain human *agency*, including actions taken to learn or use additional languages. But agency can, in a complementary or dialectical way, also influence and potentially change social structures (Deters, Gao, Miller, & Vitanova, 2015; Duff, 2012; Duff & Doherty, 2015; Flowerdew & Miller, 2008; McKay & Wong, 1996; Miller, 2014). And, although being outspoken and taking particular actions to further one's goals are observable forms of agency, intentional silence and resistance can be powerful, albeit sometimes misconstrued, forms of agency as well (Morita, 2004).

To give an example of how structure and agency work, and how identities can be constructed and contested in SLA, sometimes with disappointing learning outcomes, Talmy (2008) demonstrated how the category of “local ESL” (and “FOB,” i.e., “Fresh Off the Boat”) was produced and reproduced in Hawaiian ESL classrooms. The classes had a particular racialized group of students, many of whom were from other Pacific island communities. Instructional practices, Talmy reported, “[promoted] access to certain forms of learning and school experience, and [denied] it to others” (Talmy, 2008, p. 625). The curriculum and pedagogy were also closely connected with the local ESL learners’ observed agentive but oppositional behaviors, which included

leaving assigned materials “at home”; not doing homework; completing assignments that required minimal effort (e.g., worksheets) but not others (e.g., writing activities); starting class late; and finishing early. The more overt practices included bargaining for reduced requirements on classwork and extended time to complete it; refusal to participate in instructional activities; teasing students who did participate; and the often delicate negotiations with teachers that resulted . . .

Talmy, 2008, pp. 626–627

ESL instructors, in turn, accommodated to such dispositions and behaviors, lowering their expectations and demands of students, to the clear detriment of student learning, as demonstrated by students’ falling grades over the academic year. Interactions between macro-social structural properties and ideologies surrounding the English language education and schooling of immigrant youth such as these and students’ agency (and teachers’ acquiescence or complicity) were very evident in the classroom discourse Talmy examined.

Sex, Gender, and Sexuality

A learner’s sex, gender, or gender identity (much like ethnicity or race) is not, in itself, a good predictor of SLA. But learners of some languages may find that their ethnicity and gender interact in certain ways in relation to the target language. Male (especially working class) British, Australian, and Canadian anglophone teenagers may be reluctant to learn French, for example, not because they lack the capacity or even opportunity to learn it well, but because of how they view the language, its speakers, and culture, and how they believe being a speaker of French might position them in terms of their masculinities (Bradshaw, 2007; Carr & Pauwels, 2006; Kissau, 2006; Kissau & Turnbull, 2008; Teutsch-Dwyer, 2001). What is more, Kissau (2007) reported in his study of Canadian adolescent boys that they received less encouragement to study L2 French from their teachers, parents, guidance counselors, and peers than their female counterparts received.

Gender-based (and thus social) dimensions of SLA may also stem from folk beliefs such as “girls are better language learners” (Pomerantz, 2008). Such beliefs may contribute to some boys’ and men’s aversion to L2 education, but may also be related to the aforementioned theme of perceived and performed masculinities in what may be seen to be a highly feminized academic field and profession. For example, in a Spanish FL conversation class at a US university for which students received participation grades, Pomerantz (2008) observed that “[o]ver and over again, females were cited as preferred [classroom] partners because of their expertise in Spanish and willingness to participate in classroom activities. Males were positioned as less linguistically competent and unwilling to contribute” (p. 9). For example, Pomerantz described a second-year university student named Jim, who had

Patricia A. Duff

graduated from an elite private high school, had a positive disposition toward learning Spanish, and planned to major in it and study abroad for a year. Yet his in-class behaviors were highly inconsistent with that profile:

One possibility, as evidenced repeatedly through Jim's performance over the course of the semester, was to position oneself as a bad language user/bad language student by actively resisting classroom norms. In this way, a male student could avoid any threats to his gender identity. [. . .] In stark contrast to his punctual and well-prepared classmates, Jim often arrived late and his contributions to class discussion varied tremendously from day to day.

Pomerantz, 2008, p. 11

Jim's behaviors could be interpreted as his not wanting to curry favour with the teacher or to play the role of model language student. Another student, Ravi, aligned himself with other males in the class through his frequent joking, helping him also manage impressions of himself as not being too serious a student.

Study abroad (SA) research with university students has, similarly, revealed how, quite independent of a learner's motivation or aptitude to learn another language, race, gender, sexual orientation, social class, and/or ethnicity may affect opportunities to practice using their L2 outside of classrooms and possibly inside them as well. Indeed, reports of harassment or unwanted attentions stemming from social differences and from local cultural expectations are not uncommon in SA research (e.g., Polanyi, 1995; Talburt & Stewart, 1999). As a female African American student on a SA sojourn in Spain who received constant, negative, sexualized attention from local men lamented: "When they make commentaries to me I feel that they're taking advantage of me being different and not having command of the language. And I don't like it" (Talburt & Stewart, 1999, p. 170). Her white SA classmates then realized that their whiteness (i.e., white privilege) had mostly spared them such scrutiny and offence.

In some SA contexts, female L2 learners' experiences of SLA may be impeded not by members of the host culture or institution, as in the previous example, but by male compatriots from their home countries studying alongside them in the same language classrooms. Song (2016) conducted ethnographic research on Saudi female language students in the Southeastern US in mixed-sex, ethnically diverse classrooms that included Saudi males. She found that the female Saudi students, in the presence of Saudi males in these classrooms, were constrained by Saudi sociocultural pressures to be modest and not speak. As a result, the women deferred to the men and other classmates in interactions. Therefore, accommodations to Saudi gender-based cultural and religious sensibilities (i.e., chauvinism) mitigated against the women's full participation in classroom language use for reasons that were likely not well understood by their non-Saudi teachers or classmates and which denied them opportunities to practice English freely, the purpose of their sojourn abroad.

Peer monitoring, policing, or underperforming of in-class behaviors along gender lines is not just present among adolescent and adult language learners, however. Three studies of (im)migrant children learning English in early elementary school grades illustrate this. In Willett's (1995) study, a young male ESL learner was teased and alienated by his male classmates in a US classroom for sitting with girls at the front of the class, as required by the teacher, in order to receive additional ESL support. In another study, Toohey (2000) illustrated how a Punjabi girl in a Canadian class was also ill-treated by her classmates and misdiagnosed by her teachers as learning disabled. And Jinkerson (2011) and Mökkönen (2012), in two related articles based on the same dissertation research by the author (whose

name changed between publications), illustrated how one of the (non-Finnish) girls in her study in an English-immersion program in Finland admonished her classmates—and especially the boys—for their frequent use of Finnish instead of L2 English. She critiqued other aspects of peer behavior deemed to be inappropriate as well. In doing so, she mimicked and aligned herself with the classroom language policies of the teacher. She thereby performed the identity of “good student” (i.e., good girl) who was helping to socialize her unruly (male) classmates into proper “English-only” classroom conduct.

In addition to gender, race, and ethnicity, social class can mediate access to L2 learning experiences and to successful social inclusion and participation. In Kinginger’s (2004) study of study abroad, an American student named Alice learning French in the US, then Quebec and France, was researched over a 4-year period. Her social profile was quite distinct from that of typical middle-class students seeking to pursue SA to study French. Alice was working/lower class, with a peripatetic single mother, interrupted prior formal education, and an upbringing generally devoid of normal material comforts or even necessities. She had started working in her teens (e.g., as a waitress, exotic dancer, nanny, housekeeper, and motel receptionist), which involved many hours per week, much of it while she lived in temporary shelters or lodgings. Alice was not the typical French major that SA programs normally recruit or that textbooks cater to, according to Kinginger. Despite her unusual upbringing, however, and in contrast to the research on working-class male learners of French cited earlier, Alice had a very idealistic and romanticized view of French culture and its intellectual sophistication. Yet she “struggle[d] both with the language itself and for access to participation in social interaction” (p. 229)—further exacerbated by her lack of money, a French university system she did not understand, and mostly reluctant francophone interlocutors initially. Over time, however, Alice learned how to make strategic choices within informal francophone networks that ultimately proved very beneficial.

Thus, some learners, like Alice, may be particularly drawn to SLA precisely because of their beliefs about how being a speaker of that language might enable them to take up different identity positions. Research in Japan, for example, has shown how English education in postsecondary classrooms appeals—often quite explicitly—to the romanticized or idealistic desires (*akogare*) of young women seeking identity positions and options unavailable to them in mainstream Japanese society, which English conversation classes, and overseas travel and residence are expected to provide (Kobayashi, 2002; Kubota, 2011; Piller & Takahashi, 2006; Takahashi, 2013). In-class conversation topics and textbooks or readings also contribute or respond to *akogare* with their depictions of exciting lifestyles and relationships in the anglophone West.

Social dynamics, (mis)perceptions, or biases in relation to students’ social identities, such as those discussed earlier, may impact learners’ abilities and future trajectories; that is, whether they will reach advanced levels of proficiency and academic pursuits, or will be limited to a narrower range of future possibilities. This issue of misperceptions of learners’ backgrounds and aspirations has been observed in work with immigrant women, who were expected to take up low-skill employment or mainly domestic roles (Auerbach & Burgess, 1985; Menard-Warwick, 2008). The immigrant lower-class males described in Talmy (2008), similarly, were not envisioned to have promising academic futures. Students’ and teachers’ accommodations to such assumptions often leads to precisely that outcome of underachievement.

Research has also shown how social categories often used in SLA are far from monolithic, whether based on ethnicity or nationality (e.g., *Japanese* or *Mexican*) and sex (e.g., *female*, *male*) or institutional status as a language learner (e.g., *English language learner*, or *international university student*). Categories may be performed in completely different

Patricia A. Duff

ways and with quite distinct outcomes—even by the same students across different concurrent courses (Morita, 2004; Zappa-Hollman & Duff, 2015). For example, in Morita’s (2004) study, six “Japanese female international students” studying at a Canadian university had very different profiles from one another, in terms of their observed participation in their English-medium content courses. But even the same student participated differently in her different courses, remaining quiet in some, and more active in others, depending on a number of factors. Similarly, the set of three Mexican international university students in Zappa-Hollman and Duff’s (2015) study at a Canadian university each had unique albeit sometimes overlapping social networks (which the authors called “individual networks of practice”), which facilitated or impeded their linguistic and academic learning and achievement in their courses. Naturally, SLA scholars understand the heterogeneity represented by any social category. However, it is still necessary to better understand the conditions under which such students may be able to learn and use English to their full potential and one way of doing this is to understand the range of experiences learners in one putative category may have.

Finally, although not yet researched in SLA to the extent that sex and gender has been, research on how sexuality factors into classroom discourse and learning has been conducted by several scholars, most prominently Nelson (e.g., 2009). The argument is that, as with other types of social difference, marginalization and a lack of positive role models (in instruction or in teaching materials) can limit their opportunities to excel in SLA.

Ethnolinguistic Community Affiliation and SLA

The identification of language learners with a primary cultural group, and the ethnolinguistic vitality and identity of that group vis-à-vis the target-language group, has been examined for decades in applied linguistics (see reviews in Duff, 2012; Trofimovich & Turševa, 2015). However, an interesting context for current SLA research that challenges this dichotomy (i.e., notions of people’s first *versus* target languages, or cultures) is found in heritage-language (HL) learning communities, where the primary (L1) and target language may be one and the same (or at least closely related, such as another dialect of Chinese; Duff, Liu, & Li, 2017).

Recent research has revealed how perceptions of one’s status as HL learners can position learners in many possible ways: for example, as not needing formal instruction and thus being ineligible for language instruction in credit-bearing programs; as being proficient HL speakers, and thus given more opportunities to demonstrate their knowledge as linguistic role models in class; or as not being proficient enough in the HL or in the preferred (standard) variety of the language, even when possibly quite proficient (as in Abdi’s, 2011 Spanish HL study in Canada referred to earlier; see also Duff et al., 2017, for a discussion of learners of Chinese as an HL in Canada and the UK). Learners may choose *not* to learn their HL (e.g., Spanish or Mandarin), especially when young, precisely to differentiate themselves from immigrant populations—including their own families—and to become better integrated in their local peer cultures; they may opt to learn French or Spanish instead when they are older (He, 2010). These histories, desires, or aversions to learning their HL or aligning themselves with a wider community of speakers of their HL or one related to it often become manifest in classroom discourse and interaction as well.

Related research has illustrated the tensions and dilemmas experienced by immigrant learners in relation to the multiple communities they are part of and the “tug of war” they may experience when pulled between them by social pressures. Kim and Duff (2012), for example, described the travails of Generation 1.5 Korean students in Canada. Generation 1.5

is a social category that, like many others, represents a wide range of experiences. Typically in the North American context, Generation 1.5 students immigrate from, say, Korea, while in elementary or secondary school. Thus, they were native speakers of Korean but then have received a significant part of their academic education in English post-immigration. (They are considered “1.5” because they resemble their parents, Generation 1.0, in some respects, especially if they are highly proficient in their L1 and L1 literacy prior to immigration; but the earlier they immigrate, the more the students begin to resemble Generation 2.0, Canadian- or US-born students, especially in terms of oral English and cultural affinities). In the Kim and Duff study, participants disclosed their dilemmas about socializing with English-speaking groups, on the one hand, versus with their local/transnational Korean-speaking communities, on the other. Greater involvement with one group (and language) had negative apparent effects on their engagement with the other, according to participants. These competing and often vacillating affiliations also had a reported impact on the students’ English language development and social standing. Although not part of that particular study, it has often been reported that in-class social configurations for pair or group may reflect some of these tensions and the social pressures (and often desires) to work with same-L1 classmates rather than with students from other backgrounds (or vice versa; e.g., Duff, 2002).

Social Class

Social class was mentioned earlier in relation to gendered positionalities (e.g., as working class male or female language students). Social class is a construct that has mostly been overlooked in SLA but is now being examined much more closely (e.g., Block, 2007, 2014, 2015; Kanno & Vandrick, 2014; Vandrick, 2014). SLA studies with migrant workers in the 1970s and 1980s typically described their lack of prior education and opportunities for classroom learning but focused primarily on linguistic features in their L2s, such as basic utterance structures and how they developed cross-linguistically and over time (e.g., Klein & Purdue, 1992). More recently, classroom studies involving refugee populations have examined aspects of students’ lack of prior literacy in any language in relation to their SLA; illiteracy is often an artifact of, or proxy for, lower socioeconomic class, serial displacement, and lack of educational opportunities, especially for women with children (e.g., Bigelow, 2010; Bigelow & Watson, 2011). Now researchers are carefully examining learners’ experiences based on their socioeconomic and not just linguistic histories and trajectories.

Social class is commonly discussed in terms of lower, middle, and upper classes, and subdistinctions among them, but these descriptors are often inadequate ways of theorizing class in contemporary society. For example, (im)migrants often experience significantly reduced socioeconomic status and opportunities postmigration, or as students, in comparison with their former pre-(im)migration status, depending on their prior credentials, language proficiency, and other factors. Also relevant may be their occupations, educational attainment, possession of material resources, such as property, particular types of housing (and neighborhoods inhabited), technological or electronic tools, types of clothing or accouterments, art, vehicles, and other aspects of lifestyle, including various forms of consumption (Block, 2015). Such markers of relative wealth or capital may become quite apparent within language learning communities and classrooms (e.g., when students are asked to describe places they have traveled for leisure purposes, or other topics often featured in language textbooks and thus class discussion, or even by the way students dress and comport themselves). Social and economic capital (and symbolic capital) may also be connected with learners’ sense of entitlement and agency (or lack thereof).

Social class is therefore a relevant albeit undertheorized factor in SLA because of differences in the kinds of beliefs, dispositions, resources, and past experiences students may bring to their language learning. Students with ample financial resources, furthermore, may avail themselves of extensive extracurricular tutoring, private classes, travel, study abroad, multimedia experiences, and forms of consumption that others may simply not have access to but that are known to contribute to language learning. Conversely, if students must engage in a great deal of paid or unpaid work outside of class, they may have less time and energy to devote to their language studies, as in Alice's case, reported by Kinginger (2004) and discussed earlier. Yet, increasing one's L2 proficiency can lead to greater social access, mobility, and possibilities.

Although discussions of social class in education—including SLA—have often focused on learners with relatively few resources and with histories characterized by difficulty and deprivation, a new line of research, situated within discourses of transnationalism, flexible citizenship, and global capital (Duff, 2015; Darwin & Norton, 2014), examines those engaged in SLA who come from highly privileged social positions. Vandrick (2011), for example, describes a new “global elite” of English language learners who have lived in three or more countries, often in the pursuit of high-quality education. Such learners may manifest elements of entitlement, hybridity, and cosmopolitanism, without feeling particularly invested or rooted in any one country.

Darvin and Norton (2014) illustrate differences that class and family background can have in the life of immigrant students in English-medium school environments in Canada. They contrast the experiences of two immigrant teenagers, Ayrtton and John, both originally from the Philippines, but whose educational trajectories were distinctly different. Whereas Ayrtton came from a wealthy, entrepreneurial family, with university-educated parents, and spoke English at home (e.g., with his stay-at-home mother), John's family was more dispersed and fragmented; his mother (a midwife) and sister worked long hours in Canada and were often not at home to supervise John or his studies. In addition, John reportedly spoke an accented, nonprestigious variety of English (from his perspective), and used Filipino at home and with friends. Unlike Ayrtton, who was in an honors English program at an elite private school, John attended an inner-city public school with a large proportion of immigrants, and lacked the means for extracurricular tuition in English or other subjects. Darwin and Norton (2014) reported that John “has little opportunity to build a larger social network where he can strengthen his English skills and enter into wider conversations about social and cultural opportunities in Canadian society” (p. 115). Furthermore, they noted,

[h]is circle of friends remains resolutely local, and a great majority of them are Filipino, with whom he speaks in his mother tongue. In this peer network, or field, he has valued cultural capital. However, his relative lack of progress in English compromises his opportunities for the future.

Darvin & Norton, 2014, p. 115

Shin (2012, 2014) described some of the tensions that arise when international high school students from relatively privileged social backgrounds (e.g., middle class Koreans, in her study) interact with local Canadians, or with less wealthy Korean students from an earlier generation of immigration, or with students from other backgrounds. She described how the newcomers often looked down on these other individuals because of their perceived lack of sophistication, cosmopolitanism, and familiarity with coveted Korean popular culture. The newcomers' strong sense of entitlement and privilege in some aspects of their lives was

nonetheless often undermined by their still-developing proficiency in English and by their academic standing. Their socioeducational experiences as English language learners were thus complicated, as Shin (2014) illustrated with the experience of one grade 12 student:

Yu-ri, who was a 12th-grader in a Toronto high school at the time of the study, had studied in New Zealand before she moved to Toronto. At the school she attended in New Zealand, Yu-ri was hurt by racial slurs such as “yellow monkey” and by her White classmates who mocked her “accented” Asian English (see e.g., Lippi-Green, 1997); thus, she did not speak up in class. Even in Toronto, making friends with White Canadian students was difficult for her. She did not feel that the Korean immigrant students were welcoming either, therefore, she mostly socialized with other [study abroad] students.

pp. 100–101

A final example related to social class and SLA involves a similar demographic of international high school students (Asian students with considerable personal resources studying abroad). Such international students are now actively recruited by schools and universities in Canada and other anglophone countries seeking external sources of funding to meet their budget needs. Deschambault’s (2015) study focused on four such international students in a British Columbia high school and how they were positioned as relatively affluent English language learners but with very different histories, current contexts, and dispositions toward ESL. He examined their performance in their ESL classes and on English language and academic tests, and interviewed them and their teachers over the course of a year. The students desperately wanted to “get out of ESL” and into mainstream credit-based academic instruction with non-“ESL” students, in part because of the stigma of “ESL” and in part because ESL courses typically do not count, in terms of credits, toward high school matriculation. Their status as international fee-paying students rather than immigrant students also positioned them in ways they considered disadvantageous within the school. However, teachers seemed to know relatively little about their complicated personal circumstances and histories; furthermore, the school’s questionable means of assessing students linguistically and using relatively unchallenging (and infantilizing) classroom activities often stymied the students’ goals and wishes, at a great personal cost to them and financial cost to their families.

Other Social Dimensions of Instructed SLA

In the 1990s, the role of *participation* in learning (and participation *as* learning) became foregrounded, based on sociocultural theories related to communities of practice, situated learning, and language socialization (e.g., Lave & Wenger, 1991; McKay & Wong, 1996; Miller & Zuengler, 2011; Morita, 2004; Norton, 2001; Swain & Deters, 2007; Zappa-Hollman & Duff, 2015). Examining types of participation in SLA contexts, both inside and outside of class (e.g., in online discussion groups for classes), is therefore important. What becomes clear, however, is that participation, like identities, is very much subject to negotiation based on power differences and alliances among some members but not others (Morita, 2004). Willingness or desire to participate and sufficient L2 proficiency do not ensure inclusion, opportunities to speak, or validation within a group.

Additional social factors that may be relevant in the context of classroom activities include: working with same-L1 or different-L1 partners, or from other cultural backgrounds whose L1 may be the same or different. But as Kobayashi (2003, 2004) has shown with respect to group (L2 English) oral presentations, other social roles—even with homogeneous

Patricia A. Duff

Japanese-L1 learners of L2 English in the same SA program—may influence task-based in-class and out-of-class work and learning. How are groups formed? Who takes the lead in initiating interpretations of, and responses to, the task? How are others' roles delegated, taken up, and ultimately performed? What sorts of interpersonal factors in collaborative learning activities facilitate or impede task accomplishments? How is agency manifested within tasks? How is *intersubjectivity* (agreement, consensus) reached? And how do group members attempt to coalesce and perhaps differentiate themselves from other groups? His observation of group behaviors both in and out of class as students did task planning and rehearsals, and interviews with members of groups and instructors, shed light on some of these important questions (Kobayashi, 2003, 2004).

Other studies of classroom interaction have shown linguistic acts such as responding to a teacher's question or agreeing with another student, or having a teacher elaborate favourably upon one's contributions, can be conducive to greater participation and SLA (Atkinson, 2014; Duff, 2002). Conversely, ignoring, rebutting, or calling out another's contribution (through mockery), is evidence of disalignment or disaffiliation with the other, which can be demoralizing and counterproductive.

Empirical Evidence

Empirical evidence for the impact of social dimensions and differences among language learners, such as research themes and findings reported earlier, typically comes from case studies, auto-ethnographies, narrative inquiry, interview-based studies, and conversation analysis (drawing on, e.g., *membership categorization analysis*, in which people's affiliations with particular groups—e.g., “us vs. them”—become evident through labels, pronouns, and other referential devices they use). Ethnographic studies of classroom interactions are also common, typically employing some discourse analysis as well (e.g., Duff, 2002). Some studies use a combination of all or many of these qualitative, interpretive, and sometimes critical approaches.

Survey questionnaires are another way of ascertaining learners' perceptions of the impact of social status and social difference and have long been used in social-psychological research on attitudes and motivation toward languages, toward language learning, and toward particular ethnolinguistic groups associated with the language (Dörnyei, 2010; Duff, 2012). Matched guise techniques have also been used, when speakers' true ethnolinguistic identities are masked and listeners make judgments on the basis of the speakers' perceived identities (e.g., with bilinguals speaking in one language or another but possibly judged as *attractive*, *intelligent*, and so on, in one language but not when speaking the other, based on deep social and linguistic biases). Sometimes quantitative approaches (tests, questionnaires) are effectively combined with qualitative approaches in order to document not only changing perceptions of social dimensions of learning but also learners' actual SLA development (e.g., Kinginger, 2008).

Pedagogical Implications

The impact of social dimensions and differences in SLA can be dramatic although not necessarily visible to onlookers. Social factors can contribute to withdrawal from language study, seclusion, deep disappointment, and an early return to students' home countries during study abroad (e.g., Kinginger, 2009); conversely, they may give some learners confidence, and a sense of authority, entitlement, and agency to express themselves in their growing L2 networks. As noted earlier, learners' *social* characteristics (e.g., identities, roles, ethnic

heritage), in and of themselves, should have no direct bearing on students' SLA success in terms of their inherent *ability* or *capacity* to learn. However, if learners are viewed unfavorably, demeaned, or ostracized because of real or perceived social differences, and they are not included in or given access to meaningful and rewarding social interactions, SLA experiences and outcomes will inevitably suffer.

Some pedagogical implications are shown in the following Teaching Tips box. Most importantly, teachers need to understand their students' backgrounds and sensitivities surrounding their histories and circumstances—not simply based on impressions, stereotypes, or assumptions. In addition, teachers need to strive to maximize *all* students' opportunities to participate meaningfully and safely in language education activities through well-designed and well-monitored activities, diverse texts and topics, and varied participation structures and formats.

Teaching Tips

- Get to know your students, their backgrounds, goals, communities, and social networks.
- Don't make assumptions about students based on perceived social categories or appearances.
- Become aware of your own (mis)conceptions about issues of race, class, gender, and other social variables or areas of difference.
- Find ways of drawing upon students' backgrounds, interests, and expertise without positioning them as cultural showpieces or authorities in ways they might not appreciate (particularly as minority group members).
- Give students some degree of choice over content and in-class groupings.
- Be principled in your use of grouping strategies and closely monitor the interpersonal/social dynamics at play (inside and outside class); for example, if some students tend to monopolize discussion while others are silent, devise distinct roles for students to play that will give each a unique contribution to make.
- Give students opportunities to express but also play with "voice" and "identities"—to explore other positionalities and perspectives. This can be achieved by allowing them to take on different personae in oral and written activities and find suitable linguistic means to express perspectives from their own and others' standpoints. Such activities also expand their sociolinguistic repertoires.
- Examine stereotypical portrayals of language, identity, and social roles in media and textbook materials used in courses. Pay attention to the kinds of people (or categories, classes, experiences) that are both included in and left out of in such materials.
- Understand that phonological "accent" may index aspects of students' histories that they are proud of and that, if easily comprehended, need not be problematized.
- Confront occurrences of social exclusion, hostility, or indifference toward others both during in-class and out-of-class interactions.
- Understand and try to optimize students' social networks and opportunities to engage in language activities both inside and outside class. Consider alternative ways in which students can participate meaningfully; for example, by posting comments online allowing students time to compose contributions and not only through spontaneous speech; or through the use of i-clickers (interactive response systems/tools) during large-class discussions or lectures.
- Vary participation formats, by using pair and small-group work (and different combinations of students) and not just large-class formats.

Patricia A. Duff

When teachers or textbook activities state: “Let’s have a girl and a boy do this role play: You [male student] be the businessman, and you [female student] the office assistant,” the language experiences, utterances, and roles students are invited to take up are not of equal status, and perpetuate certain cultural and gender-based stereotypes. Similarly, identifying a heritage language learner in a heterogeneous Mandarin class and saying: “You’re Chinese, you’ll be able to read the Mandarin text out loud for us,” might deny another student, not of apparent Chinese heritage, the opportunity to be a model speaker; it also makes assumptions about the proficiency of heritage learners or their willingness to be defined and showcased as such. In Abdi’s (2011) research cited earlier, for example, in high school Spanish-L2 classrooms, a teacher’s invitations of this sort to particular students were problematic; in doing so, she privileged those students and not others. Furthermore, asking students to talk about “their homelands and cultures” in the L2 positions students as not being locally born or raised, which can be deeply disturbing to them, particularly for Generation 1.5 students (Talmy, 2008). Teachers need to be aware of such dynamics and misgivings.

Future Directions

This chapter has underscored social dimensions and differences relevant to SLA. An increasing range of studies—particularly longitudinal, ethnographic case studies across different learning, linguistic and geographical contexts—will help applied linguists and language educators better understand the complex sociological forces at work in classroom interactions and learning. Furthermore, as theories and constructs evolve to better address new forms of transnationalism and engagements in language study, SLA will develop new means of conceptualizing, conducting, interpreting, and representing empirical research in this area.

Note

1. The category of age here is understood to be both a biological factor—linked to relative neural plasticity or maturation—and a social category, on which basis there may be increased or reduced opportunities to engage in language use or to be included in language-mediated activity based on (perceived) age differences among participants. For example, middle-aged or older sojourners in a study-abroad program may have very different opportunities, social networks, and experiences than students in their early twenties.

References

- Abdi, K. (2011). “She really only speaks English”: Positioning, language ideology, and heritage language learners. *Canadian Modern Language Review*, 67(2), 161–190.
- Atkinson, D. (Ed.). (2011). *Alternative approaches to second language acquisition*. New York: Routledge.
- Atkinson, D. (2014). Language learning in mindbodyworld: A sociocognitive approach to second language acquisition. *Language Teaching*, 47, 467–483.
- Auerbach, E., & Burgess, D. (1985). The hidden curriculum of survival ESL. *TESOL Quarterly*, 19(3), 475–495.
- Batstone, R. (Ed.). (2010). *Sociocognitive perspectives on language use and language learning* (pp. 3–23). Oxford: Oxford University Press.
- Bigelow, M. (2010). *Mogadishu on the Mississippi: Language, racialized identity, and education in a new land*. New York: Wiley-Blackwell.

- Bigelow, M., & Watson, J. (2011). Educational level and L2 learning. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 461–475). New York: Routledge/Taylor Francis.
- Block, D. (2003). *The social turn in second language acquisition*. Edinburgh: Edinburgh University Press.
- Block, D. (2007). *Second language identities*. London: Continuum.
- Block, D. (2014). *Social class in applied linguistics*. London: Routledge.
- Block, D. (2015). Social class in applied linguistics. *Annual Review of Applied Linguistics*, 35, 1–19.
- Block, D., & Corona, V. (2014). Exploring class-based intersectionality. *Language, Culture and Curriculum*, 27(1), 27–42.
- Bradshaw, J.M. (2007). Are second language classrooms gendered? In H. Marriott, T. Moore, & R. Spence-Brown (Eds.), *Learning discourses and the discourses of learning* (pp. 1–17). Clayton, Victoria: Monash University ePress.
- Carr, J., & Pauwels, A. (2006). *Boys and foreign language learning: Real boys don't do languages*. New York: Palgrave Macmillan.
- Darvin, R., & Norton, B. (2014). Social class, identity, and migrant students. *Journal of Language, Identity, and Education*, 13(2), 111–117.
- Deschambault, R. (2015). *Economizing education: Fee-paying ESL students in a public high school* (PhD dissertation). University of British Columbia. Retrieved from <https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0220586>
- Deters, P., Gao, X., Miller, E. R., & Vitanova, G. (Eds.). (2015). *Interdisciplinary approaches to theorizing and analyzing agency and second language learning*. Bristol: Multilingual Matters.
- Dörnyei, Z. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). London: Routledge.
- Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, 100-S, 19–47.
- Duff, P. (2002). The discursive co-construction of knowledge, identity, and difference: An ethnography of communication in the high school mainstream. *Applied Linguistics*, 23, 289–322.
- Duff, P. (2012). Identity, agency, and SLA. In A. Mackey & S. Gass (Eds.), *Handbook of second language acquisition* (pp. 410–426). London: Routledge.
- Duff, P. (2015). Transnationalism, multilingualism, and identity. *Annual Review of Applied Linguistics*, 35, 57–80.
- Duff, P., Anderson, T., Ilnyckj, R., VanGaya, E., Wang, R., & Yates, E. (2013). *Learning Chinese: Linguistic, sociocultural, and narrative perspectives*. Berlin/New York: De Gruyter.
- Duff, P., & Doherty, L. (2015). Examining agency in (second) language socialization research. In P. Deters, X. Gao, E. Miller, & G. Vitanova (Eds.), *Interdisciplinary approaches to theorizing and analyzing agency and second language learning* (pp. 54–72). Bristol: Multilingual Matters.
- Duff, P., Liu, Y., & Li, D. (2017). Chinese heritage language learning: Negotiating identities, ideologies, and institutionalization (pp. 409–422). In O. Kagan, M. Carreira, & C. Chik (Eds.), *Routledge handbook on heritage language education*. New York: Routledge.
- Duff, P., & Surtees, V. (in press). In A. Burns & J. Richards (Eds.), *Cambridge guide to learning a second language*. Cambridge: Cambridge University Press.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Erard, M. (2012). *Babel no more*. New York: Free Press.
- Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal*, 81(3), 285–300.
- Flowerdew, J., & Miller, L. (2008). Social structure and individual agency in second language learning: Evidence from three life histories. *Critical Inquiry in Language Studies*, 5(4), 201–224.
- He, A. W. (2010). The heart of heritage: Sociocultural dimensions of heritage language acquisition. *Annual Review of Applied Linguistics*, 30, 66–82.
- Jinkerson, A. C. (2011). Interpreting and managing a monolingual norm in an English speaking class in Finland: When first and second graders contest the norm. *Apples—Journal of Applied Language Studies*, 5(1), 27–48.

Patricia A. Duff

- Kanno, Y., & Vandrick, S. (Eds.). (2014). Social class in language learning and teaching: Forum section. *Journal of Language, Identity, and Education*, 13(2), 85–123.
- Kim, J., & Duff, P. (2012). The language socialization and identity negotiations of Generation 1.5 Korean-Canadian university students. *TESL Canada Journal*, 29(6), 81–102.
- Kinginger, C. (2004). Alice doesn't live here anymore: Foreign language learning and identity construction. In A. Pavlenko & A. Blackledge (Eds.), *Negotiation of identities in multilingual contexts* (pp. 219–242). Clevedon: Multilingual Matters.
- Kinginger, C. (2008). Language learning in study abroad: Case studies of Americans in France. *Modern Language Journal*, 92(S1), 1–124.
- Kinginger, C. (2009). *Language learning and study abroad*. London: Palgrave Macmillan.
- Kissau, S. (2006). Gender differences in motivation to learn French. *Canadian Modern Language Review*, 62, 401–422.
- Kissau, S. (2007). Is what's good for the goose good for the gander? The case of male and female encouragement to study French. *Foreign Language Annals*, 40, 365–378.
- Kissau, S., & Turnbull, M. (2008). Boys and French as a second language: A research agenda for greater understanding. *Canadian Journal of Applied Linguistics*, 11(3), 151–170.
- Klein, W., & Purdue, C. (1992). *Utterance structure: Developing grammars again*. Amsterdam: John Benjamins.
- Kobayashi, M. (2003). The role of peer support in ESL students' accomplishments of oral academic tasks. *Canadian Modern Language Review*, 59(3), 337–369.
- Kobayashi, M. (2004). *A sociocultural study of second language tasks: Activity, agency, and language socialization* (Doctoral dissertation). University of British Columbia. Retrieved from <https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0078243>
- Kobayashi, Y. (2002). The role of gender in foreign language learning attitudes: Female students' attitudes towards English learning. *Gender and Education*, 14, 181–197.
- Kubota, R. (2011). Learning a foreign language as leisure and consumption: Enjoyment, desire, and the business of *eikaiwa*. *International Journal of Bilingual Education and Bilingualism*, 14(4), 473–488.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States*. London: Routledge.
- McCall, L. (2005). The complexity of intersectionality. *Signs*, 30(3), 1771–1800.
- McKay, S. L., & Wong, S.-L. C. (1996). Multiple discourses, multiple identities: Investment and agency in second-language learning among Chinese adolescent immigrant students. *Harvard Educational Review*, 66(3), 577–609.
- Menard-Warwick, J. (2005). Both a fiction and an existential fact: Theorizing identity in second language acquisition and literacy studies. *Linguistics and Education*, 16, 253–274.
- Menard-Warwick, J. (2008). “Because she made beds. Every day”: Social positioning, classroom discourse, and language learning. *Applied Linguistics*, 29, 267–289.
- Miller, E. R. (2014). *The language of adult immigrants: Agency in the making*. Bristol: Multilingual Matters.
- Miller, E. R., & Zuengler, J. (2011). Negotiating access to learning through resistance to classroom practice. *Modern Language Journal*, 95, 130–147.
- Mökkönen, A. C. (2012). Social organization through teacher-talk: Subteaching, socialization and the normative use of language in a multilingual primary class. *Linguistics and Education*, 23(3), 310–322.
- Morita, N. (2004). Negotiating participation and identity in second language academic communities. *TESOL Quarterly*, 39, 573–603.
- Nelson, C. D. (2009). *Sexual identities in English language education: Classroom conversations*. New York: Routledge.

- Norton, B. (2001). Non-participation, imagined communities and the language classroom. In M. Breen (Ed.), *Learner contributions to language learning: New directions in research* (pp. 159–171). Harlow: Pearson Education.
- Norton, B. (2013). *Identity and language learning: Extending the conversation* (2nd ed.). Bristol, UK: Multilingual Matters.
- Norton, B., & Toohey, K. (2001). Changing perspectives on good language learners. *TESOL Quarterly*, 35(2), 307–322.
- Piller, I., & Takahashi, K. (2006). A passion for English: Desire and the language market. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression and representation* (pp. 59–83). Clevedon: Multilingual Matters.
- Polanyi, L. (1995). Language learning and living abroad: Stories from the field. In B.F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 271–291). Philadelphia: John Benjamins.
- Pomerantz, A. (2008). “Señora, boys just cannot speak Spanish”: Negotiating masculine identities in a foreign language classroom. *Working Papers in Educational Linguistics*, 23(2), 1–23.
- Shin, H. (2012). From FOB to COOL: Transnational migrant students in Toronto and the styling of global linguistic capital. *Journal of Sociolinguistics*, 16, 184–200.
- Shin, H. (2014). Social class, habitus, and language learning: The case of Korean early study-abroad students. *Journal Language, Identity, and Education*, 13, 99–103.
- Song, J. (2016, April). *Gender, identity, and language socialization among Saudi female students in the U.S.* Paper presented at the annual meeting of the American Association for Applied Linguistics, Orlando, FL.
- Swain, M., & Deters, P. (2007). “New” mainstream SLA theory: Expanded and enriched. *Modern Language Journal*, 91, 820–836.
- Takahashi, K. (2013). *Language learning, gender and desire: Japanese women on the move*. Bristol: Multilingual Matters.
- Talbot, S., & Stewart, M. (1999). What’s the subject of study abroad?: Race, gender, and “living culture.” *Modern Language Journal*, 83, 163–175.
- Talmy, S. (2008). The cultural productions of the ESL student at Tradewinds High: Contingency, multidirectionality, and identity in L2 socialization. *Applied Linguistics*, 29, 619–644.
- Teutsch-Dwyer, M. (2001). (Re)constructing masculinity in a new linguistic reality. In A. Pavlenko, A. Blackledge, I. Piller, & M. Teutsch-Dwyer (Eds.), *Multilingualism, second language acquisition, and gender* (pp. 175–198). New York: Mouton de Gruyter.
- Toohey, K. (2000). *Learning English at school: Identity, social relations and classroom practice*. Clevedon: Multilingual Matters.
- Trofimovich, P., & Turuševa, L. (2015). Ethnic identity and second language learning. *Annual Review of Applied Linguistics*, 35, 234–252.
- Vandrick, S. (2011). Students of the new global elite. *TESOL Quarterly*, 45, 160–169.
- Vandrick, S. (2014). The role of social class in English language education. *Journal of Language, Identity, and Education*, 13(2), 85–91.
- Willett, J. (1995). Becoming first graders in an L2: An ethnographic study of L2 socialization. *TESOL Quarterly*, 29(3), 473–503.
- Zappa-Hollman, S., & Duff, P. (2015). Academic English socialization through individual networks of practice. *TESOL Quarterly*, 49, 333–368.

Cognitive Differences and ISLA

Shaofeng Li

Background

Whereas Chapter 21 concerns the influence of the social aspects of individual differences in ISLA, this chapter focuses on the cognitive dimensions, particularly language aptitude and working memory, which have instigated a large body of research in the past few decades (see Li, 2015; Linck, Osthus, Koeth, & Bunting, 2014). Interest in learners' cognitive differences is motivated by their explanatory power in accounting for variability in L2 learning and the valuable implications the research findings have for practitioners regarding how to tailor instruction to achieve maximal instructional effects. In the following, I introduce the basics of the two constructs, elaborate the theories and controversies, synthesize the research findings, and discuss ways to incorporate the research findings into L2 pedagogy. By way of clarification, while there has been a call to consider working memory as a component of language aptitude, due to the lack of research mapping the associations between the two cognitive variables and to the existence of parallel streams of research on them, they are dealt with separately in this chapter.

Language Aptitude

Traditionally language aptitude refers to a set of cognitive abilities, including phonetic coding ability, language analytic ability, and rote memory, which are predictive of learning rate. This conglomerate of abilities is postulated to be (1) the initial state of readiness for foreign language learning, (2) relatively stable or not subject to training, learning experience, or environmental factors, (3) distinct from other individual difference variables such as motivation and anxiety, and (4) domain specific in the sense that it is exclusive to learning a foreign language and therefore different from intelligence or abilities for learning other academic subjects (Carroll, 1981). Among these characteristics, (1) and (2) await further empirical verification, and (3) and (4) are supported by Li's meta-analysis (2016) that aggregated the correlations between aptitude and other individual differences reported in primary studies. The meta-analysis revealed that language aptitude is uncorrelated with motivation and negatively correlated with anxiety, and that it has a large overlap, but is not isomorphic, with intelligence.

The history of language aptitude research dates back to the 1950s when large-scale initiatives were undertaken to develop and validate aptitude tests for the purpose of selecting qualified learners for state-funded language programs in the US and Canada where learners were expected to master a foreign language through short-term intensive training. The most influential aptitude test that has dominated the research is the Modern Language Aptitude Test (MLAT) (Carroll & Sapon, 1959), which consists of five subtests that measure the three components of aptitude—phonetic coding, language analytic ability, and memory. Other tests that were subsequently developed include the PLAB (Pimsleur Language Aptitude Battery)—a test for high school learners (Grades 7–12); the DLAB (Defence Language Aptitude Battery), which targets high-aptitude learners (Petersen & Al-Haik, 1976); VORD (“vord” is the word for “word” in the artificial language used in the test), which is intended for the learning of challenging languages (Child, 1998); the CANAL-F (Cognitive Ability for Novelty in Acquisition of Language [Foreign]), which aims to test abilities to “cope with novelty and ambiguity” (Grigorenko, Sternberg, & Ehrman, 2000, p. 392); and the LLAMA (Meara, 2005), a free computerized test modelled on the MLAT. Note that to date the existing aptitude measures have not been cross-validated and therefore the extent to which they measure the same construct is uncertain.

The traditional conceptualization of language aptitude and the way it is measured have been criticized on a number of grounds. It has been argued that the abilities measured via existing tests are important only for learning the formal aspects of language and for learning language as discrete items, and that they do not account for how the pragmatic aspects of a language are learned and how learning happens in communicative tasks (Skehan, 2012). It is also argued that these abilities are important only for preliminary but not advanced L2 learning. To identify abilities for advanced learning, Linck et al. (2013) developed an aptitude battery called the Hi-LAB, which consists of 12 measures of seven cognitive abilities, and found measures of phonological short-term memory, implicit learning, and rote memory to be significant predictors of high attainment. One feature that stands out about the Hi-LAB is the inclusion of six measures of the executive and storage functions of working memory, which demonstrates the researchers’ emphasis on the importance of this cognitive device in advanced learning. In the following, I provide the background information about working memory and how it relates to traditional aptitude.

Working Memory

Working memory refers to the ability to simultaneously store and manipulate incoming stimuli (Baddeley, 2007). It consists of a central executive, a phonological loop, a visuospatial sketchpad, and an episodic buffer. The central executive has no storage capacity and is responsible for shifting attention between meaning and form and between information retrieval and task performance, inhibiting irrelevant information, and coordinating between the subsystems (Juffs & Harrington, 2012). The phonological loop is a space where verbal information is stored and rehearsed. The sketchpad deals with visuospatial information such as images, shapes, and locations. Second language (L2) learning, as Baddeley (2015) pointed out, relates only to the verbal and attentional aspects of working memory, which explains why there has been very little research on the visuospatial sketchpad. Baddeley’s initial model consists of only the domain-general central executive and the two domain-specific subsystems. Later Baddeley recognized the need for a component that serves as a bridge between the two subsystems and the central executive, linking short-term memory with long-term memory and integrating discrete items into larger units, hence the episodic buffer. The episodic buffer is a relatively recent addition to the model and has not been extensively investigated.

Working memory has been measured in two ways—through (1) simple tasks that tap only the storage component and (2) complex tasks that gauge both the storage and processing components. A simple task requires learners to remember unrelated items such as digits and nonwords, while a complex task typically consists of two parts: one that requires the learner to conduct some semantic, syntactic, or mathematical processing and one that requires the learner to remember an element of the item in question. For example, in a typical reading or listening span test, the learner reads or hears groups of sentences that vary in the number of included sentences, judges the plausibility of each sentence (whether the meaning makes sense or whether it is grammatical) while remembering the final word of each sentence, and recalls the sentence-final words at the end of each group. The measurement of working memory by means of these two different task types leads to, or rather represents, two distinguishable streams of research. One, led by Baddeley and associates, has mainly investigated the role of the phonological loop—measured through the digit or word repetition tasks—in studies of L1 vocabulary learning. The other, initiated by Daneman and Carpenter (1980), has investigated the importance of working memory using complex tasks in studies of higher order abilities such as reading comprehension. Wen (2015) referred to these two streams of research as the British camp versus the North American camp. In this chapter, the term “working memory” is used to refer to the concept of short-term memory, and “phonological short-term memory” to the storage component (the phonological loop), which is measured through simple tasks. Where it is necessary, the term “complex working memory” is used to denote both the storage and processing components.

Key Concepts

Language Aptitude

- Components of traditional language aptitude: phonetic coding, language analytic ability, and rote memory.
- Characteristics of traditional aptitude
 - predictive of learning rate
 - stable
 - important for initial language learning
 - distinct from motivation and anxiety but correlated with intelligence.

Working Memory

- Architecture of working memory
 - central executive: The boss of working memory that monitors the system
 - phonological loop: The inner ear that stores and rehearses verbal information
 - visuospatial sketchpad: The inner eye that deals with images, shapes, and locations
 - episodic buffer: The interim buffer that integrates information into larger units and accesses long-term memory.
- Measures of working memory
 - complex tasks: Listening span, reading span, operation span, backward digit span
 - simple tasks: Word span, nonword span, forward digit span.

In general working memory has been found to be distinct from other cognitive variables. For example, in a meta-analysis on the associations between working memory and intelligence, Ackerman, Beier, and Boyle (2005) found that the correlation was $r = .48$ for complex working memory measures and $r = .35$ for simple measures, the magnitudes of the correlations indicating that working memory and intelligence are not identical constructs. Working memory has also been found to be separate from language aptitude although there has been a call to incorporate it as a component of aptitude. For example, Hummel (2009) found no significant correlations between phonological short-term memory (the phonological loop) and aptitude as measured by the French version of the MLAT. Roehr and Gánem-Gutiérrez (2009) found that working memory gauged through L1 and L2 reading span tests loaded on different factors than components of aptitude measured using the MLAT. However, this study reported a significant correlation between the learners' scores of L1 reading span and their composite MLAT scores. Therefore, it is necessary to look more closely at the associations between working memory and language aptitude.

Current Issues

Current research on cognitive differences in instructed L2 learning draws on different methodological paradigms: a predictive approach and an interactional approach (Li, 2015). The purpose of predictive research is to identify variables that are important for the final learning outcome, regardless of instruction type and learning context. Underlying such a perspective is a preference for an eclectic approach to language instruction (Carroll, 1963; Scrivener, 2005) and the assumption that there is no need to tailor instruction to accommodate individual differences. In a typical predictive study, two sets of scores are obtained, one for a predictor variable such as aptitude (Sparks, Patton, Ganschow, & Humbach, 2011) or working memory (Harrington & Sawyer, 1992), and one for a criterion variable such as general L2 proficiency or some specific aspect of learning such as listening comprehension. Analyses of a correlational nature (Pearson's correlation, multiple regression analysis, etc.) are then conducted to determine whether the ID (individual difference) variable is a significant predictor. This constitutes a static, product-oriented approach.

In the predictive paradigm, issues of current interest relating to language aptitude include the associations between overall aptitude measured through whole test batteries and aptitude components measured through subtests on the one hand, and outcome measures for general L2 proficiency and specific aspects of learning on the other. Also of interest are whether traditional aptitude is predictive of only initial learning but not advanced proficiency, and whether it is implicated only in traditional foreign language classes but not in more meaning-oriented instruction such as immersion and communicative language teaching. For working memory, of primary interest is whether the two types of short-term memory, tested through complex and simple tasks, are correlated with L2 outcomes and whether they have differential effects on different aspects of learning.

The interactional approach draws heavily on the ATI (aptitude-treatment-interaction) model (Cronbach & Snow, 1977; Dance & Neufeld, 1988; Snow, 1991) from educational psychology. In this approach, ID variables are viewed as dynamic constructs that interact with instruction type, and the effectiveness of an instructional task type depends on whether there is a fit between task type and the learner's cognitive profile. Interactional studies are experimental and are conducted to investigate the comparative effects of different treatment types and how these effects are related to scores for one or more cognitive variables (e.g., Sheen, 2007). The instructional treatments are characterized by consistent

manipulation of variables and use of focused tasks and tests that target one or several particular linguistic structures. In this respect, the interactional approach differs from the predictive approach, where learners are not tested on their knowledge or use of particular structures.

Within the interactional paradigm, researchers have been interested to know whether the role of aptitude varies across different learning conditions such as different types of corrective feedback, deductive and inductive instruction, and implicit and explicit instruction, due to the different processing demands imposed on the learner. With respect to working memory, in addition to examining whether it is associated with the learning that happens in different instructional treatments, one area that has aroused some interest but is insufficiently researched is whether working memory affects learners' speech performance in different ways in tasks that vary along the procedural (e.g., with or without planning) and conceptual (e.g., with or without reasoning demand) dimensions of complexity.

Empirical Evidence

Language Aptitude

Predictive Research

The primary objective of predictive aptitude research is to see whether aptitude can be used to forecast how well, compared with peers, one can master a foreign language within a given time period (Carroll, 1981). It has been found that aptitude measured by whole test batteries such as the MLAT is a strong predictor of general L2 proficiency measured by course grades (Carroll & Sapon, 2002) or standardized proficiency tests (e.g., Sparks, Patton, Ganschow, & Humbach, 2009). However, the meta-analysis by Li (2015) found that high school foreign language learners were more likely to draw on aptitude than university foreign language learners, suggesting that aptitude is more important at initial stages of L2 learning, given that high school students are generally beginners and are normally less proficient than university students.

Two primary studies have investigated proficiency as an independent variable and produced mixed results. Winke (2005) reported a study on the relationship between aptitude and L2 Chinese achievement among first semester learners at Georgetown University and advanced learners at the Defence Language Institute who underwent 63 weeks of intensive training. Winke reported that aptitude was correlated only with the achievement scores of the beginning learners but not those of the advanced learners. In another study, Hummel (2009) found that aptitude was predictive of the proficiency scores of a group of advanced ESL learners, but when the learners were divided into high and low proficiency using the median score as the cutoff point, aptitude was no longer a significant predictor for either proficiency level. The robustness of the results of the two studies is compromised by some methodological limitations such as the use of different proficiency tests for the two groups of learners in Winke's study, the relatively low reliability of the aptitude test (a French version of the MLAT, Cronbach's $\alpha = .55$), and the inclusion of only three components of L2 achievement (vocabulary, grammar, and reading) in the proficiency test in Hummel's study. More research is warranted on whether aptitude has differential effects on different phases of L2 development or whether different sets of abilities are implicated at different stages of learning.

Li (2016) detected two other noteworthy patterns regarding the predictive validity of aptitude as a global construct. One is that the original MLAT (the English version)

showed a stronger predictive validity than the MLAT designed for other languages such as the Hungarian version (HUNLAT; Sáfár & Kormos, 2008), or the French version (TALV; Hummel, 2009). Thus the disparate findings about the associations between aptitude and criterion variables in some previous studies may be partly due to the different aptitude tests used in the studies (Granena & Long, 2013). Second, while overall aptitude has been found to be a significant predictor of general proficiency and some specific aspects of learning such as grammar (Bialystok & Fröhlich, 1978), listening (Keitges, 1986), reading (Ehrman, 1998), and speaking (Sparks, Patton, Ganschow, & Humbach, 2012), it has not been a significant predictor of vocabulary learning (Winke, 2005) and writing (Sáfár & Kormos, 2008).

Although traditional aptitude tests such as MLAT, PLAB, and VORD were validated by correlating overall aptitude scores and criterion variables such as course grades, empirical studies have examined the associations between different aptitude components and L2 learning. First, it has been found that language analytic ability is a strong predictor of grammar learning (e.g., DeKeyser, 1993; Gardner & Lambert, 1965), which is of no surprise given the hypothesized link between this aptitude component and the learning of the morphosyntactic aspects of an L2. However, one issue that needs to be resolved is whether language analytic ability is drawn on in the learning of implicit or explicit linguistic knowledge. Although a distinction between the two knowledge types is not made in previous studies, it would seem that the item-based written grammar tests in most of the studies encouraged or allowed the use of explicit knowledge, which traditional aptitude measures are sensitive to (Granena & Long, 2013). Second, phonetic coding has been found to be strongly correlated with vocabulary learning (e.g., Sparks et al., 2011), suggesting that the ability for bottom-up processing of unfamiliar sounds is critical to learning new words in a foreign language. Third, among all the aptitude components, rote memory is the least predictive of L2 learning, including vocabulary learning, which is somewhat surprising in light of its putative importance in memorizing word translations (Li, 2015). The weak predictive validity of rote memory constitutes a justification for including alternative memory measures such as working memory as an aptitude component. However, there needs to be more research on the theoretical and empirical links between traditional aptitude and working memory. If both working memory and rote memory are significant predictors of L2 achievements, as Linck et al. (2013) reported in their validation study on the Hi-LAB aptitude test, or if each explains a unique portion of the variance of SLA, it is necessary to include both as components of language aptitude.

Most predictive studies are conducted with foreign language classes (in settings where the target language is not the language of the community) that are heavily form oriented and that may favour the abilities measured by traditional aptitude tests, which begs the question of whether aptitude is relevant in more meaning-oriented contexts. Ranta (2002) reported that language analytic ability measured by means of an L1 metalinguistic test was a significant predictor of the learning outcomes of communicative classes. However, it is uncertain whether a measure of L1 metalinguistic knowledge is a valid test of language analytic ability. Harley and Hart (1997, 2002) reported an interaction between aptitude and age in an immersion setting: younger starters relied more on memory while older starters were more likely to draw on language analytic ability. These two studies provide preliminary evidence that aptitude is also correlated with the learning that happens in meaning-focused instruction and that learners of different age groups draw on different aptitude components.

In short, the predictive aptitude research has found that:

1. Composite aptitude scores are strongly and consistently predictive of L2 learning achievements except for vocabulary learning and L2 writing.
2. Aptitude is likely more associated with initial than higher levels of learning.
3. Language analytic ability is a significant predictor of grammar learning.
4. Phonetic coding ability is important for vocabulary learning.
5. Rote memory is a consistently weaker predictor of general proficiency and specific aspects of learning than other aptitude components.
6. Aptitude seems also relevant in meaning-oriented instruction, not only in form-oriented instruction.
7. Child L2 learners tend to rely more on memory and adult learners more on analytic ability, although this claim needs to be tested in further research.

Interactional Research

Interactional aptitude research falls into three categories, examining the mediating effects of aptitude in (1) deductive and inductive instruction, (2) implicit and explicit treatments, and (3) different types of corrective feedback. Within the first category, one oft-cited study is Erlam (2005), which investigated the relationships between two aptitude components—language analytic ability and phonetic coding—and three instructional types in the learning of French object pronouns by L1 English-speaking students in a New Zealand secondary school. It was found that language analytic ability was significantly correlated with the effects of inductive instruction and structured input but not deductive instruction. In a study exploring whether high- and low-aptitude learners benefitted differently from inductive and deductive instruction, Hwu and Sun (2012, 2014) included three aptitude components: memory for text (ability to memorize grammar rules), analytic ability and rote memory, which were treated as one construct. The results showed that deductive instruction was significantly more effective for the low-aptitude learners than inductive instruction, but the reverse was true for high-aptitude learners, although the result was not statistically significant.

The results of Hwu et al.'s study seem at odds with those of Hauptman (1971), where high-aptitude learners benefitted more from a situational approach where grammar explanation was provided deductively than a structural approach where rules were taught inductively. However, in Hauptman's study, the two treatment types differed in other ways in addition to how the grammar rules were presented. The structural approach entailed sequencing the linguistic material in order of increasing difficulty and the heavy use of drills and mechanical practice, whereas in the situational approach, materials were not sequenced linguistically and practice happened mainly through role play. Thus it would seem that linguistic materials in the situational class were less structured and therefore potentially required higher abilities. Also, although grammar was taught inductively, the kind of exercises (substitution, blank-filling, etc.) likely favoured grammar learning and thus negated the role of differences in aptitude.

These studies seem to show that (1) the role of aptitude is less important in deductive instruction where more external support (in the form of rule explanation) is available, and therefore deductive instruction favours low-aptitude learners who need more external support; and (2) high-aptitude learners benefit more from inductive instruction that pushes them to exploit their cognitive resources. These inferences are in line with Snow's (1991)

argument that high-structured tasks help less able learners, while low-structure tasks are best for high-ability learners. However, as can be seen, the studies on deductive and inductive instruction were conducted in very different ways, which makes it difficult to draw firm conclusions. For example, in Erlam's (2005) study, the inductive group was never provided with explicit grammar explanation, whereas in Hwu et al.'s study (according to a footnote in the 2012 report), the inductive instruction included metalinguistic feedback that is equivalent to rule explanation. Also, aptitude was operationalized differently in the three studies—as analytic ability and phonetic coding in Erlam (2005), as a cluster of three components in Hwu, Wei, and Sun (2014), and as a composite construct measured by a whole aptitude battery in Hauptman (1971).

A second line of research concerns whether language aptitude is sensitive only to explicit learning conditions. According to Krashen (1981, p. 158), “what is considered second or foreign language aptitude may be directly related to conscious learning”—a hypothesis that seems to have been confirmed by several studies (Carpenter, 2008; de Graaff, 1997; Robinson, 1997, 2002). These studies share some methodological features: they all included implicit and explicit computer-delivered treatments. These studies show that aptitude is likely to be more correlated with the learning that happens under conditions where learners engage in conscious processing of linguistic forms and less likely in implicit and incidental learning conditions that do not direct learners' attention to forms or require learners to process meaning only. Although de Graaff's study (1997) showed significant correlations between aptitude and the effects of the implicit treatment, the treatment included form-focused activities that raised learners' awareness of the linguistic targets and therefore was not entirely implicit.

Finally, a number of studies have investigated the role of language analytic ability in different feedback conditions. Sheen (2007) found that this cognitive ability was predictive of only the effects of metalinguistic feedback, not those of recasts. Yilmaz (2013) reported that meta-linguistic feedback was more effective than recasts only when learners had high analytic ability. These two studies seem to indicate that, similar to the findings of studies on implicit and explicit instructional treatments, aptitude is more relevant in explicit feedback conditions. However, two studies that examined computer-mediated feedback reported that aptitude was also important in implicit conditions when no feedback (Sachs, 2010) and recasts (Trofimovich, Ammar, & Gatbonton, 2007) were provided. The significant effects of aptitude on learning under the implicit feedback conditions in the two studies might be due to the possibility that the instructional treatments are not entirely implicit.

Li (2013a, 2013b) reported a complicated interaction between aptitude, feedback type, and the nature of the linguistic target. It was found that in the learning of Chinese classifiers, analytic ability was correlated with the effects of recasts but not those of metalinguistic feedback. In the learning of the Chinese aspect marker *-le*, however, the reverse was true. Li speculated that this was because in the case of classifiers—a syntactically transparent structure—the provision of metalinguistic explanation neutralized the role of aptitude. In the recast condition where metalinguistic information about classifiers was unavailable, the influence of language analytic ability became evident. However, the aspect marker involves complicated linguistic projections and required the learners to use their analytic ability to process the rule explanation available in the metalinguistic condition. When the rule explanation was unavailable, as in the recast condition, it was beyond the learners' ability to induce the rule of the complicated structure using their own cognitive resources, which explains why recasts were ineffective in the learning of this target structure.

The following is a summary of the findings of the interactional aptitude studies:

1. Aptitude is less important in deductive instruction than in inductive instruction.
2. Low-aptitude learners benefit more from deductive instruction.
3. High-aptitude learners benefit more from inductive instruction.
4. Explicit instruction is more likely to implicate aptitude than implicit instruction.

Working Memory

Predictive Research

For an overview of the growing body of research on the predictive power of working memory in ISLA, a good starting point is Linck et al.'s meta-analysis (2014), which aggregated the results of 79 studies involving 3,707 learners. In this study, working memory measures were coded as simple or complex based on task type, as L1 or L2 in terms of language of performance, and as verbal (e.g., listening span, word span) or nonverbal (e.g., operation span, digit span) depending on whether the task involved the processing of linguistic information. Outcome measures were divided into comprehension (reading comprehension, grammar test, etc.) and production (cloze test, translation, global proficiency test, etc.), and into processing and proficiency, with the former involving online linguistic processing such as picture description and the latter assessments of L2 "knowledge and more general language abilities" (p. 866) such as vocabulary and narrative abilities. The overall correlation between all working memory measures and outcome measures was $r = .25$, suggesting that working memory has a significant, albeit weak, correlation with L2 learning. Furthermore, simple tasks were less predictive than complex tasks of L2 learning, $r = .17$ versus $r = .27$, especially for measures of proficiency, and verbal measures showed stronger associations with L2 outcomes than nonverbal measures.

The meta-analysis also found that L1 measures, especially those of complex working memory, showed weaker correlations with L2 outcomes than L2 measures, suggesting that learners' performance on L2 working memory tests is related to their L2 proficiency. Other sources of evidence also suggest that learners' working memory performance might be influenced by their L2 proficiency and that L1 and L2 measures may tap different constructs. For example, Jongejan, Verhoeven, and Siegel (2007) found that on an English working memory test, L1 English children's scores were higher than those of their ESL peers. Similarly, Walter (2004) found that L1 French learners' working memory scores were substantially higher on a French (L1) test than on an English (L2) test. Also, in studies that reported correlations between L1 and L2 working memory scores (Alptekin & Ercetin, 2010; Geva & Ryan, 1993; Harrington & Sawyer, 1992; Juffs, 2005), most are in the range of .3–.5, and only in rare cases (e.g., $r = .84$ in Osaka & Osaka, 1992) are higher correlations observed. To address the potential effect of L2 proficiency, one approach is to "employ L1 measures to provide a purer estimate of WM abilities" (Linck et al., 2014, p. 872) and another is to include a measure of proficiency and identify the unique contribution of working memory after the influence of proficiency has been accounted for.

Although Linck et al.'s study showed that simple memory tasks (i.e., phonological short-term memory) were less predictive than complex tasks, it is premature to conclude that phonological short-term memory is less important because the two types of memory may play complementary roles (Wen, 2015) and facilitate different aspects and stages of L2 learning. In the following, I provide a synthesis of the research on the two types of short-term

memory—phonological short-term memory and complex working memory—in terms of their roles in different aspects of L2 learning.

Vocabulary

Phonological short-term memory has been found to be relevant to vocabulary learning (Engel de Abreu & Gathercole, 2012; Hummel, 2009; Speciale, Ellis, & Bywater, 2004) due to its putative function as a device for learning new words (Baddeley, 2015). There is preliminary evidence that phonological short-term memory is important only for the initial stages of vocabulary learning while at more advanced stages previous vocabulary knowledge or long-term memory becomes more important (Cheung, 1996; Masoura & Gathercole, 2005). The limited research on complex working memory revealed that it is not predictive of child L2 vocabulary learning (Jean & Geva, 2009; Jongejan et al., 2007) but it has positive effects on adult vocabulary learning (Kempe, Brooks, & Christman, 2009; Martin & Ellis, 2012).

Grammar

There are two possible ways phonological short-term memory impacts on grammar learning. One is indirectly through vocabulary learning, that is, words and formulaic sequences learned through phonological short-term memory provide data for linguistic analysis and rule-learning (Williams, 2012). The other way is directly through memorizing and extracting the rules governing sequences of morphemes. The hypotheses have been confirmed in Martin and Ellis's (2012) study where phonological short-term memory was found to have a direct effect on grammar learning and an indirect effect via vocabulary learning. A few other studies also reported significant correlations between phonological short-term memory and grammar learning (Daneman & Case, 1981; French & O'Brien, 2008; Hummel, 2009). In a recent study by Serafini and Sanz (2016), phonological short-term memory was found to be a predictor of the L2 Spanish grammatical knowledge of beginning and intermediate learners, but not that of advanced learners.

Complex working memory has also been found to be significantly correlated with L2 grammar learning (e.g., Engel de Abreu & Gathercole, 2012). When both phonological short-term memory and complex working memory are examined, it is often the latter that shows stronger predictive validity (Harrington & Sawyer, 1992; Martin & Ellis, 2012), which is attributed to the processing element it involves. However, somewhat surprisingly, complex working memory has been found to have no effect on online syntactic processing. Juffs and Harrington (2012) suggested that this might be because individual differences in working memory are eclipsed by L1 processing habits in L2 sentence processing.

Reading

Whereas phonological short-term memory is predictive of vocabulary and grammar learning, it has not been implicated in reading comprehension (Geva & Ryan, 1993; Harrington & Sawyer, 1992; Hummel, 2009). However, similar to L1 reading comprehension (see Daneman & Merikle, 1996, for a meta-analysis), L2 reading comprehension has shown consistent, positive correlations with complex working memory (Fontanini & Tomitch, 2009; Harrington & Sawyer, 1992; Payne, Kalibatseva, & Jungers, 2009). However, one general theme that has emerged is that L2, but not L1, working memory measures are predictive of L2 reading comprehension (Alptekin & Ercetin, 2010; Geva & Ryan, 1993; Harrington &

Shaofeng Li

Sawyer, 1992; Miyake & Friedman, 1998), which, once again, suggests an effect of L2 proficiency on L2 working memory performance. Finally, Walter (2004) found that complex working memory was significantly correlated only with low-intermediate ESL learners' reading comprehension ability, measured through a gapped summary completion task, but not upper-intermediate learners' reading ability.

Speaking

There has been limited predictive research on the associations between the two types of short-term memory and L2 speaking. Two studies have reported positive correlations between phonological short-term memory and the development of L2 oral proficiency, namely gains between two time points (O'Brien, Segalowitz, Freed, & Collentine, 2007; Payne & Whitney, 2002). Positive correlations were found between complex working memory and oral fluency (Fehringer & Fry, 2007) as well as overall oral competence (Kormos & Sáfár, 2008). Payne and Whitney (2002) failed to find a significant correlation between complex working memory and the *development* of oral proficiency. The relevance of phonological short-term memory to improvement in oral abilities and of complex working memory to oral performance remains to be further investigated.

The findings of predictive working memory research are summarized as follows:

1. Complex working memory is more predictive of L2 achievements than phonological short-term memory.
2. Verbal working memory measures are more predictive than nonverbal measures.
3. L2 measures are more predictive than L1 measures, suggesting an impact of L2 proficiency on the research findings.
4. Phonological short-term memory is a significant predictor of vocabulary learning while the role of complex working memory in vocabulary learning is uncertain.
5. Phonological short-term memory is more important for vocabulary learning at the beginning stage, while at more advanced stages, long-term memory or learners' existing vocabulary knowledge takes over as the more dominant factor for vocabulary development.
6. Complex working memory appears to be a stronger predictor of grammar learning than phonological short-term memory, although both are significant predictors. Complex working memory is not important for online syntactic processing.
7. Complex working memory is a stronger predictor of reading comprehension than phonological short-term memory. However, it appears that the predictive power of working memory is evident only when it is measured in learners' L2 but not when it is measured in their L1.
8. Phonological short-term memory facilitates the *development* of oral proficiency while working memory is important for oral *performance*.

Interactional Research

Interactional working memory studies fall into two categories: those examining the mediating role of working memory in affecting the *learning* that results from interactional feedback and those exploring the effects of working memory on task *performance*. Most of the studies investigated complex working memory rather than phonological short-term memory, because of the assumed importance of the former in online information processing, which

characterizes the instructional treatments or task prompts of the interactional studies. Feedback studies are couched in the Interaction Hypothesis, which emphasizes the importance of attending to linguistic forms in meaning-focused communicative tasks, necessitating a heavy reliance on working memory. When receiving feedback, the learner must mobilize his/her working memory resources to tune in to the information contained in the feedback, maintain it in accessible state, and retrieve information from long-term memory to process the information.

Most studies on L2 oral task performance are based on the Limited Attention Capacity Hypothesis (LACH) (Skehan, 2014) and the Cognition Hypothesis (Robinson, 2011). The LACH, drawing on Levelt's (1989) model of speech production, states that speech production undergoes three phases: conceptualizing the message, formulating the language representation or finding the linguistic forms for the message, and articulating the message. The three stages involve controlled processing, which is effortful and conscious and poses heavy demands on working memory resources. The Cognition Hypothesis holds that task complexity can be increased along two groups of variables—resource-directing variables (e.g., +/- reasoning demands) that direct learners' cognitive resources to the notions and corresponding linguistic resources, and resource-dispersing variables (+/- planning) pertaining to the procedural aspects of tasks. Increasing task complexity along resource-directing dimensions diverts learners' working memory resources to advanced linguistic structures and leads to the use of complex language. Increasing task complexity along resource-dispersing variables depletes learners' cognitive resources and has detrimental effects on task performance. According to the Cognition Hypothesis, the role of working memory is more evident in complex tasks, which are more cognitively demanding than simple tasks.

Corrective Feedback

Several studies have examined the role of working memory in noticing the corrective force of recasts and in L2 development. Mackey, Philp, Egi, Fujii, and Tatsumi (2002) discovered that learners with high working memory reported more noticing of recasts provided on L1 Japanese speakers' errors relating to English question formation in dyadic interaction. In terms of L2 development, learners with smaller working memory capacities showed more immediate gains and those with high working memory abilities demonstrated more delayed gains. Kim, Payant, and Pearson (2015) also reported that working memory was a significant predictor of ESL learners' noticing of recasts and development in question formation. The study also found that more learners with high working memory in the complex task (with higher reasoning demand) advanced to higher stages of question formation than in the simple task. Révész (2012) investigated the influence of complex working memory and phonological short-term memory on the effects of recasts in the learning of the English past progressive tense by Hungarian ESL learners. Treatment effects were measured using one oral task and two written tests. Significant correlations were found between complex working memory and gains on the written tests, and between phonological short-term memory and gains on the oral test. Révész argued that complex working memory facilitates the learning of explicit/declarative knowledge while phonological short-term memory is more conducive to the acquisition of implicit/procedural knowledge.

A few studies (Goo, 2012; Li, 2013a, 2013b; Yilmaz, 2013) probed the interaction between working memory and feedback type—implicit feedback in the form of recasts versus explicit feedback operationalized as metalinguistic feedback in Goo's and Li's studies

and as explicit correction in Yilmaz's. Working memory was not correlated with the effects of implicit feedback in Li's and Yilmaz's studies, but it was in Goo's study. However, the reverse was found for explicit feedback: while Li and Yilmaz found significant effects for working memory, Goo did not. Furthermore, Li found a negative association between working memory and the effects of explicit feedback in the learning of a complicated linguistic structure—the Chinese aspect marker *-le*, which may occur in a postverbal or sentence-final position and that is subject to multiple interpretations.

It is difficult to disentangle the conflicting results of the foregoing studies, which may result from the methodological inconsistencies between them in terms of the target structure, instructional setting, learners' proficiency level, treatment task, and so on. However, one commonality between the studies is that the researchers explained the presence or absence of a significant effect of working memory by resorting to its noticing function, that is, the impact of working memory surfaces when the treatment condition orients the learner's attention to the information contained in corrective feedback. This explanation is confirmed by the preliminary results of an ongoing meta-analysis (Li, in progress) that shows stronger associations between working memory and the effects of explicit instructional treatments in comparison with implicit treatments—consistent with what is found about language aptitude.

Task Performance

Studies on task performance have focused on two resource-dispersing variables—task planning and task structure—and one resource-directing variable—with or without reasoning demand. With regard to task planning, Ellis (2005) distinguished pretask planning (also called strategic planning) and within-task planning (or online planning). Pretask planning allows the learner to think about the language and content prior to task performance but imposes a time pressure for task performance. Within-task planning allows the learner to perform the task without time pressure and encourages the learner to think about the content and language during, rather than before, task performance. In many task-based studies, within-task planning is either not controlled or there is a lack of information on whether or not it is controlled. Ahmadian (2012) is one of the few studies examining the role of working memory in unpressured within-task planning. It is reported that working memory was significantly correlated with accuracy and fluency but not complexity of a group of Iranian ESL learners' narrative production. Another study by Li and Fu (in press) sought to ascertain whether working memory plays different roles under strategic and unpressured within-task planning conditions. Working memory was found to be significantly correlated with accuracy and fluency in the within-task planning condition, but it was not correlated with the performance of the strategic planning group. The authors argued that the role of working memory is evident during unpressured performance because it affords opportunities for learners to monitor their production. Such opportunities are unavailable during pressured performance, which explains why there is a lack of significant effects in the strategic planning condition. The absence of significant correlations for the strategic planners may also be attributable to the eased burden on message conceptualization as a result of the opportunity for pretask planning.

Kormos and Trebits (2011) undertook a study to see whether structured and unstructured tasks drew on working memory in different ways. In the structured task, the learners told a narrative based on a set of cartoon pictures sequenced in a coherent order, while in the unstructured task the learners had to invent a story based on unrelated pictures. In both tasks,

learners were allowed 2 minutes to plan before task performance, but it is unclear whether a time limit was imposed for task performance. Somewhat surprisingly, it is the structured task rather than the unstructured task that showed significant correlations with working memory. The authors speculated that this is because although the structured task alleviated the burden on the conceptualization in terms of content planning, it increased the demand on the formulation in that the learners had to select linguistic items to match the prescribed content. This study also found an uneven relationship between working memory and task performance. For example, learners with larger working memory capacities performed better in terms of clause length, but they produced fewer subordinate clauses.

Manipulating task complexity along a different dimension, Crespo (2011) conducted a study on the interface between working memory and task complexity operationalized as with or without reasoning demand. Adult L1 Spanish EFL learners performed two versions of the same decision-making task, the more complex version requiring learners to figure out the relationships between more elements, consider more factors when making decisions, and have access to fewer resources. Three aspects of working memory were examined: working memory as a global construct for storage and processing, phonological short-term memory, and attention control (the central executive). The results revealed that only phonological short-term memory was a significant predictor, for both simple and complex tasks. The study failed to confirm Robinson's prediction that complex tasks are more likely to draw on working memory, and it also suggested that despite the putative link between working memory and online task performance, phonological short-term memory may turn out to be a crucial factor for speech production.

To conclude, the following claims can be made based on the interactional working memory studies:

1. Working memory facilitates the noticing of recasts.
2. Working memory is implicated when learners receive corrective feedback during communication that requires them to juggle between form and meaning.
3. It is possible that working memory facilitates the learning of explicit knowledge while phonological short-term memory enhances the acquisition of implicit knowledge.
4. Working memory is drawn upon in unpressured performance but not pressured performance after pretask planning.
5. Pretask planning may ease the burden on message conceptualization and thus neutralize the adverse effect of low working memory.
6. Tasks that provide a clear structure for performance may tax learners' working memory resources to a greater extent than tasks without a clear structure—contrary to what is commonly assumed.
7. Increasing the reasoning demand of a task may not necessarily pose a greater challenge for working memory.
8. The role of phonological short-term memory in oral task performance may be of particular significance.

Pedagogical Implications

The predictive research on language aptitude and working memory shows that aptitude is a strong determinant of L2 success and therefore should be taken into account when making pedagogical decisions. Carroll and Sapon (2002), who developed the MLAT primarily for predictive purposes, made a number of recommendations on ways to use students' aptitude

scores, including selecting ideal learners for state-funded language programs or classes where one is expected to master a foreign language in a short period through intensive training; placing students with comparable aptitude levels into parallel sessions to make sure they progress at similar rates; diagnosing learning abilities to provide guidance; waive foreign language requirements; and match learner types with instructional approaches.

The recommendation for matching learner types with instructional approaches is of special importance to teachers. The assumption underlying such a recommendation is that (1) learners have different aptitude profiles, that is, one may excel in certain abilities but be poor in others and (2) in order to maximize instructional effects, there should be a fit between learners' aptitude strengths and the cognitive demands of the instruction. By way of illustration, Weshe (1981) reported a model developed by the Canadian Public Service Commission French language program where three instructional approaches were adopted: an audiovisual approach, an analytic approach, and a functional approach. The audiovisual approach focuses on dialog memorization and drills but excludes grammar explanation, translation, and use of reading and writing "in the early phases of training" (Wesche, 1981, p. 127). The latter two approaches were developed to accommodate highly analytic and memory-oriented learners respectively; the two types of learners were distinguished based on their aptitude scores and reported preferences. The analytic approach emphasized grammar instruction and use of drills and written exercises, whereas the functional approach featured meaning-oriented activities such as role play and games. Weshe reported the results of a verification project (internal report submitted to a government agency), which discovered the superior effects of matching analytic learners with the analytic approach, in comparison with the unmatched condition where analytic learners were forced to follow the audio-visual approach. The author admitted that this was not a rigorously designed experimental study because there is uncertainty over the distinctions between the three approaches and whether they were consistently implemented.

Whereas the preceding study was a longitudinal project concerning the macro aspects of the so-called aptitude-treatment interaction (Snow, 1991), interactional aptitude studies where experimental procedures are carefully designed to minimize the interference of extraneous variables are more revealing about the methods and instructional techniques teachers may employ to address learner differences. To begin with, the research shows that the role of aptitude tends to be neutralized in deductive instruction, and that low-aptitude learners benefit more from deductive instruction while high-aptitude learners more from inductive instruction. In the spirit of maximizing instructional effects for learners of different aptitude profiles and catering to the meaning-primary principle of the currently popular task-based instruction, it would seem advisable to employ inductive tasks that prompt learners to discover rules through meaning-oriented tasks and then provide explicit rule explanation in the posttask stage to accommodate low-aptitude learners who need more external assistance. Providing rule explanation followed by practice through communicative tasks, as in the deductive approach, may predispose learners to focus on linguistic form rather than to allocate primary attention to meaning (Ellis, 2003; Willis & Willis, 2007).

Second, the feedback research indicates that aptitude and working memory are more likely to be drawn on in tasks with an explicit focus on form, which disadvantages low-aptitude learners. However, because overall explicit feedback has proven more effective than implicit feedback (Ellis, Loewen, & Erlam, 2006; Li, 2010), at least in terms of short-term effects, it is advisable to make the corrective intention known to the learner when feedback is used as a form-focusing device to facilitate L2 development. One way to accommodate learners with lower aptitude levels and weaker working memory abilities is to provide

pretask instruction, a practice further buttressed by the superior effects of pretask instruction plus task-embedded feedback compared with feedback or instruction alone (Li, Ellis, & Zhu, 2016a). Alternatively, feedback can be delayed until the posttask stage when the task is over. However, there is preliminary empirical evidence and theoretical basis for the superiority of immediate feedback to delayed feedback in enhancing L2 development (Li, Ellis, & Zhu, 2016b). Furthermore, the hypothesis that providing pretask instruction or offline (delayed) feedback reduces learners' cognitive burden needs to be empirically tested.

Third, the research on working memory demonstrates that learners make heavy use of their memory resources during unpressured performance but not during pressured performance after pretask planning, which may alleviate the burden on working memory. However, the research on task planning (Li & Fu, in press; Yuan & Ellis, 2003) shows that within-task planning allows learners to monitor their production and leads to greater accuracy. Thus the best option seems to be allowing learners to plan both before and during task performance. It is also found that structured tasks that are assumed to be simpler than unstructured tasks may turn out to be more complex because in a structured task, the demand for formulating stipulated content is higher and consequently the task may be more taxing on working memory resources. One way to support learners with low working memory abilities during unstructured tasks is to provide some task-essential linguistic input in the form of key words or expressions before or during task performance (Robinson, 2007).

Finally, all recommendations regarding how to accommodate learners' cognitive differences are based on the assumption that teachers have the information about their students' cognitive profiles. There are two ways to ascertain learners' cognitive strengths or weaknesses: through subjective and objective methods. Subjective methods include asking students to self-report, such as via an interview (Weshe, 1981) or questionnaire (Granena, 2016), their cognitive propensities or preferences, and/or observing their learning behaviours in the classroom. Although the validity of self-reported information is questionable, there has been empirical evidence that shows significant correlations between learners' self-reported cognitive styles and their performance on aptitude tests (e.g., Granena, 2016). Tests of aptitude and working memory may provide more reliable information but validated tests such as the MLAT and the PLAB are not accessible to teachers. One free aptitude test that is electronically available is the LLAMA (which can be easily found through Google), which has been recently used in a number of published studies (e.g., Granena & Long, 2013). As to measures of working memory, many published articles (e.g., Hummel, 2009) provide either example items or full-length tests in the appendices, and the tests can be easily administered in class or in a computer lab.

Teaching Tips

- Example uses of aptitude scores:
 - Selecting students
 - Counselling
 - Placement
 - Diagnosing learning abilities
 - Waiving language requirements.
- Keep in mind that different approaches and methods favour learners with different cognitive strengths.

- Use instruction types that are more effective than other instruction types for all learners but adapt aspects that disadvantage learners of certain cognitive profiles.
- Use inductive tasks but provide explicit explanation at the end.
- Make feedback salient but consider pretask instruction and posttask feedback as ways of easing learners' cognitive burden.
- Allow learners to plan both before and within task performance.
- Provide linguistic support for cognitively demanding tasks.

Future Directions

Given that traditional aptitude has been found to be more important in heavily form-based instruction that is amenable to conscious learning, one promising area of research is identifying abilities that are important in implicit or unconscious learning (Granena, 2013, 2016). Also, because traditional aptitude is more correlated with preliminary stages of L2 learning, future research should probe abilities important for learning at more advanced levels. With respect to interactional research, one limitation is the inconsistency in the operationalization of instructional treatments such as inductive versus deductive and explicit versus implicit, which makes it difficult to reach unequivocal conclusions. Thus there is a need to clearly define (preferably on theoretical grounds), consistently implement and repeatedly replicate certain instructional treatments in order to obtain more robust results and make definitive claims about the associations between aptitude and the effects of different instructional treatments.

Although there has been a plethora of research on working memory, there is confusion over the construct and how it should be measured. First, there are both theoretical and empirical grounds for separating complex working memory and phonological short-term memory. Theoretically, the former refers to both the storage and processing functions while the latter to only the storage component. Empirically, the two types of working memory have been found to have differential predictive validities for L2 achievements and for different aspects of learning. Therefore, conflating the two types of short-term memory, as in some L2 studies, is not justified. Second, given the lower predictive power of nonverbal stimuli such as digit span and operation span in comparison with verbal stimuli such as word span and reading span, it is advisable to prioritize using verbal tests in the measurement of working memory in future research. Third, because of the possible influence of learners' L2 proficiency on their working memory scores, test items should not be presented in the target language; otherwise, the variance explained by learners' L2 proficiency must be accounted for. Finally, research shows that as learners move to higher proficiency levels, phonological short-term memory has diminished effects on vocabulary (Cheung, 1996) and grammar learning (Serafini & Sanz, 2016) and complex working memory shows weaker correlations with reading comprehension (Walter, 2004). However, these findings are preliminary, and there is a clear need for more research on whether working memory plays different roles at different stages of learning.

References

- Ackerman, P., Beier, M., & Boyle, M. (2005). Working memory and intelligence. *Psychological Bulletin*, 131, 30–60.
- Ahmadian, M. J. (2012). The relationship between working memory capacity and L2 oral performance under task-based careful online planning condition. *TESOL Quarterly*, 46, 165–175.

- Alptekin, C., & Ercetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, 22, 206–219.
- Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.
- Baddeley, A. (2015). Working memory in second language learning. In Z. Wen, M. Mota, & A. McNeil (Eds.), *Working memory in second language acquisition and processing* (pp. 17–28). Bristol: Multilingual Matters.
- Bialystok, E., & Fröhlich, M. (1978). Variables of classroom achievement in second language learning. *Modern Language Journal*, 62(7), 327–336.
- Carpenter, H. (2008). *A behavioral and electrophysiological investigation of different aptitudes for L2 grammar in learners equated for proficiency level* (Unpublished doctoral dissertation). Georgetown University.
- Carroll, J. (1963). A model of school learning. *Teachers College Record*, 64, 723–733.
- Carroll, J. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.
- Carroll, J., & Sapon, S. (1959). *Modern language aptitude test*. New York: The Psychological Corporation/Harcourt Brace Jovanovich.
- Carroll, J., & Sapon, S. (2002). *Manual for the MLAT*. North Bethesda, MD: Second Language Testing, Inc.
- Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary learning. *Developmental Psychology*, 32, 867–873.
- Child, J. (1998). Language aptitude testing: Learners and applications. *Applied Language Learning*, 9, 1–10.
- Crespo, M. (2011). *The effects of task complexity on L2 oral production as mediated by differences in working memory capacity* (Unpublished master's thesis). University of Barcelona.
- Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods*. New York: Irvington.
- Dance, K., Richard, W., & Neufeld, J. (1988). Aptitude-treatment interaction research in the clinical setting: A review of attempts to dispel the “patient uniformity” myth. *Psychological Bulletin*, 104, 192–213.
- Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Daneman, M., & Case, R. (1981). Syntactic form, semantic complexity, and short-term memory: Influences on children's acquisition of new linguistic structures. *Developmental Psychology*, 17, 367–378.
- Daneman, M., & Merikle, P.M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422–433.
- de Graaff, R. (1997). The experanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249–276.
- DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *Modern Language Journal*, 77, 501–514.
- Ehrman, M. (1998). The modern language aptitude test for predicting learning success and advising students. *Applied Language Learning*, 9, 31–70.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (Ed.). (2005). *Planning and task performance in a second language*. Philadelphia: John Benjamins.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–368.
- Engel de Abreu, P.M., & Gathercole, S.E. (2012). Executive and phonological processes in second-language acquisition. *Journal of Educational Psychology*, 104, 974–986.
- Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, 9, 147–171.
- Fehringer, C., & Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers: The role of working memory. *Folia Linguistica*, 41(1–2), 37–72.

- Fontanini, I., & Tomitch, L.M.B. (2009). Working memory capacity and L2 university students' comprehension of linear texts and hypertexts. *IJES*, 9(2), 1–18.
- French, L.M., & O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics*, 29, 463–487.
- Gardner, R.C., & Lambert, W.E. (1965). Language aptitude, intelligence, and second-language achievement. *Journal of Educational Psychology*, 56(4), 191–199.
- Geva, E., & Ryan, E. (1993). Linguistic and cognitive correlates of academic skills in first and second languages. *Language Learning*, 43, 5–42.
- Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in Second Language Acquisition*, 34, 445–474.
- Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63, 665–704.
- Granena, G. (2016). Cognitive aptitudes for implicit and explicit learning and information-processing styles: An individual differences study. *Applied Psycholinguistics*, 37, 577–600.
- Granena, G., & Long, M.H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311–343.
- Grigorenko, E., Sternberg, R., & Ehrman, M. (2000). A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal*, 84, 390–405.
- Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition*, 19(3), 379–400.
- Harley, B., & Hart, D. (2002). Age, aptitude, and second language learning on a bilingual exchange. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 302–330). Amsterdam/Philadelphia: John Benjamins.
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25–38.
- Hauptman, P. (1971). A structural approach vs. a situational approach to foreign-language teaching. *Language Learning*, 21, 235–244.
- Hummel, K. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics*, 30, 225–249.
- Hwu, F., & Sun, S. (2012). The aptitude-treatment interaction effects on the learning of grammar rules. *System*, 40, 505–521.
- Hwu, F., Wei, P., & Sun, S. (2014). Aptitude-treatment interaction effects on explicit rule learning: A latent growth curve analysis. *Language Teaching Research*, 18, 294–319.
- Jean, M., & Geva, E. (2009). The development of vocabulary in English as a second language and its role in predicting word recognition ability. *Applied Psycholinguistics*, 30, 153–185.
- Jongejan, W., Verhoeven, L., & Siegel, L. (2007). Predictors of reading and spelling abilities in first- and second-language learners. *Journal of Educational Psychology*, 99, 535–551.
- Juffs, A. (2005). The influence of first language on the processing of wh- movement in English as a second language. *Second Language Research*, 21, 121–151.
- Juffs, A., & Harrington, M. (2012). Aspects of working memory in L2 learning. *Language Teaching*, 44(2), 137–166.
- Keitges, D.J. (1986). *Relationships between foreign-language aptitude, attitudes and motivation, and personality traits, and the attained English language proficiency of Japanese university students* (Unpublished doctoral dissertation). University of Texas, Austin.
- Kempe, V., Brooks, P.J., & Christman, S.D. (2009). Inconsistent handedness is linked to more successful foreign language vocabulary learning. *Psychonomic Bulletin & Review*, 16(3), 480–485.
- Kim, Y., Payant, C., & Pearson, P. (2015). The intersection of task-based interaction, task complexity, and working memory. *Studies in Second Language Acquisition*, 37, 549–581.
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11(2), 261–271.

- Kormos, J., & Trebits, A. (2011). Working memory capacity and narrative task performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 267–285). Amsterdam: John Benjamins.
- Krashen, S. (1981). Aptitude and attitude in relation to second language acquisition and learning. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 155–175). Rowley, MA: Newbury House.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Li, S. (2013a). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal*, 97, 634–654.
- Li, S. (2013b). The differential roles of language analytic ability and working memory in mediating the effects of two types of feedback on the acquisition of an opaque linguistic structure. In C. Sanz & B. Lado (Eds.), *Individual differences, L2 development & language program administration: From theory to application* (pp. 32–52). Boston, MA: Cengage Learning.
- Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, 36, 385–408.
- Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition*, 38, 801–342.
- Li, S. (in progress). *The role of working memory in SLA: A meta-analysis*.
- Li, S., Ellis, R., & Zhu, Y. (2016a). Task-based versus task-supported language instruction: An experimental study. *Annual Review of Applied Linguistics*, 36, 205–229.
- Li, S., Ellis, R., & Zhu, Y. (2016b). The effects of the timing of corrective feedback on the acquisition of a new linguistic structure. *Modern Language Journal*, 100, 276–295.
- Li, S., & Fu, M. (in press). Strategic and unpressured within-task planning and their associations with working memory. *Language Teaching Research*.
- Linck, J., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, 63(3), 530–566.
- Linck, J., Osthus, P., Koeth, J. T., & Bunting, M. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21, 861–883.
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback, and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181–209). Philadelphia: John Benjamins.
- Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, 34, 179–413.
- Masoura, E. V., & Gathercole, S. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory*, 13, 422–429.
- Meara, P. (2005). *LLAMA Language aptitude tests*. Swansea: Lognostics.
- Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Lawrence Erlbaum.
- O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition*, 29, 557–582.
- Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30, 287–289.
- Payne, T. W., Kalibateva, Z., & Jungers, M. K. (2009). Does domain experience compensate for working memory capacity in second language reading comprehension? *Learning and Individual Differences*, 19, 119–123.
- Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory and interlanguage development. *CALICO Journal*, 20(1), 7–32.

- Petersen, C., & Al-Haik, A. (1976). The development of the Defence Language Aptitude Battery (DLAB). *Educational and Psychological Measurement*, 36, 369–380.
- Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 159–179). Amsterdam: John Benjamins.
- Révész, A. (2012). Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning*, 62(1), 93–132.
- Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47, 45–99.
- Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 212–266). Amsterdam/Philadelphia: John Benjamins.
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects of speech production, interaction, uptake, and perceptions of task complexity. *International Review of Applied Linguistics*, 45, 193–214.
- Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 3–38). Amsterdam: John Benjamins.
- Roehr, K., & Gánem-Gutiérrez, G. (2009). The status of metalinguistic knowledge in instructed adult L2 learning. *Language Awareness*, 18(2), 165–181.
- Sachs, R. (2010). *Individual differences and the effectiveness of visual feedback of reflexive binding in L2 Japanese* (Unpublish doctoral dissertation). Washington, DC: Georgetown University.
- Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *International Review of Applied Linguistics in Language Teaching*, 46(2), 113–136.
- Scrivener, J. (2005). *Language teaching: A guidebook for English language teachers*. Oxford: Macmillan Education.
- Serafini, E., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, 38, 607–646.
- Sheen, Y. (2007). The effects of corrective feedback, language aptitude, and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 301–322). New York: Oxford University Press.
- Skehan, P. (2012). Language aptitude. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 381–395). New York: Routledge.
- Skehan, P. (Ed.). (2014). *Processing perspectives on task performance*. Amsterdam: John Benjamins.
- Snow, R. E. (1991). Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology*, 59, 205–216.
- Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2009). Long-term relationships among early first language skills, second language aptitude, second language affect, and later second language proficiency. *Applied Psycholinguistics*, 30, 725–755.
- Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2011). Subcomponents of second-language aptitude and second-language proficiency. *Modern Language Journal*, 95(2), 253–273.
- Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Relationships among L1 print exposure and early L1 literacy skills, L2 aptitude, and L2 proficiency. *Reading and Writing*, 25, 1599–1634.
- Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, 25, 293–321.
- Trofimovich, P., Ammar, A., & Gatbonton, E. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 79–99). Oxford: Oxford University Press.
- Walter, C. (2004). Transfer of reading comprehension skills to L2 linked to mental representation of text and to L2 working memory. *Applied Linguistics*, 25, 315–339.

- Wen, Z. (2015). Working memory in second language acquisition and processing: The phonological/executive model. In Z. Wen, B. Mailce, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 41–62). Bristol: Multilingual Matters.
- Wesche, M. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 119–175). Rowley, MA: Newbury House.
- Williams, J. (2012). Working memory and SLA. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 427–441). New York: Routledge.
- Willis, D., & Willis, J. (2007). *Doing task-based instruction*. Oxford: Oxford University Press.
- Winke, P. (2005). *Individual differences in adult Chinese second language acquisition: The relationships between aptitude, memory, and strategies for language learning* (Unpublished doctoral dissertation). Georgetown University.
- Yilmaz, Y. (2013). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics*, 34, 344–368.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1–27.

Motivation in the L2 Classroom

Kata Csizér

Background

Second language (L2) motivation research is one of the most vibrant fields of applied linguistics (most recently see Csizér & Magid, 2014; Dörnyei & Ushioda, 2011; Dörnyei, MacIntyre, & Henry, 2015; Dörnyei & Ryan, 2015), with the general aim of investigating how much effort students are willing to invest into L2 learning and what might be the sources of the differences between motivated and unmotivated students (Dörnyei, 2009). The prevalence of L2 motivation research stems from the fact that motivation has long been seen as the key variable to successful L2 learning (Dörnyei & Ushioda, 2011), which has resulted in hundreds of articles published on L2 motivation and research flourishing in several distinct directions. Despite the fact that some classroom-related issues have been investigated, such as the role of the teacher or tasks in motivation (Dörnyei & Kubanyiova, 2014), empirical studies directed specifically at instructed second language acquisition (ISLA) are still relatively scarce and the differentiation between L2 motivation in instructed and naturalistic settings is usually not explored. How can we explain this relative scarcity of research output in such a vibrant field? One possible explanation is given by Crookes and Schmidt (1991), who in their classic work on the theory of motivation point out that “in informal learning, as in formal classroom learning, the basic motivational issues are the same: does the learner take advantage of opportunities for learning, persist at what is basically a difficult enterprise, and what factors facilitate such persistence?” (p. 494). A second explanation might be—and I believe most L2 motivation researchers would agree—that researchers are primarily interested in the language learner as opposed to the impact that L2 instruction may have on the learner, and therefore, most of the issues investigated in the field have concerned themselves with language learners’ characteristics, more precisely, their attitudes and dispositions, in whatever contexts they are learning the language. Third, the role of the teacher and the effect of instruction in L2 motivation research seem to be somewhat sensitive topics due to the fact that many studies point toward the demotivating roles teachers can play in the classroom (see later for details); therefore, it is possible that researchers will shy away from researching the actual effect teachers might have on (de)motivating students. Fourth, many of the investigations in L2 motivation involve large-scale samples drawn from many classrooms with the understanding that ISLA is investigated, but without going into specific details on the instruction itself (Dörnyei, Csizér, & Németh, 2006). Fifth, it can be argued that in ISLA settings the wider social context

is just as important as classroom variables; therefore, research cannot be limited by classroom environments (Candlin & Mercer, 2001). Still, I would like to argue that despite all these reasons classroom-level and instruction-related variables could and should be added to empirical L2 motivation studies. Even more so, it has been increasingly acknowledged and emphasised that contextual information in research should not only serve as a background to empirical investigations, but context should be a pivotal part of any research projects; therefore, individual difference variables, and among them L2 motivation, “enter into some interaction with the situational parameters rather than cutting across tasks and environments” (Dörnyei, 2005, p. 218).

In this chapter I set out to accomplish a number of things. First, I give a brief theoretical overview on L2 motivation looking into the extent to which the most important L2 motivational theories take ISLA into account. Second, I review empirical evidence on L2 motivation in the classroom: attitudes and the role of the teacher and the learning group will be given special emphasis. Third, pedagogical implications are discussed. Finally, I summarise possible research directions/ideas on L2 motivation and its role in ISLA.

Before providing the brief theoretical overview, it is important to provide a definition for motivation. The notion of L2 motivation is said to be difficult to define because it represents a multidimensional, complex phenomenon trying to explain human behavior (Dörnyei & Ushioda, 2011). Still, most L2 motivation researchers agree that motivation consists of a directed behavior of effort, persistence, and choice (Dörnyei & Ushioda, 2011). Choice usually refers to the fact that L2 students choose to learn, while both effort and persistence relate to the learning process itself: the amount of energy invested into language learning and how long students persevere.

As for the theoretical background of L2 motivation studies, the rich history of L2 motivation research makes it impossible to provide a detailed description of the development of this field, and therefore, it is customary to streamline various investigations into phases in order to help readers better understand the background. One such differentiation has been offered by Dörnyei (most recently Dörnyei & Ryan, 2015), who posits that L2 motivation research has had three main phases: social-psychological, cognitive-situated, and process-oriented (Dörnyei & Ryan, 2015). Despite the fact that these stages of research are usually positioned on a timeline indicating possible development in the field, for the purpose of the present chapter, I look at these phases as representing various interests in L2 motivation research, and I discuss how these main topical approaches relate to ISLA.

Key Concepts (based on Dörnyei & Ushioda, 2011)

Motivation: The amount of effort invested into a specific behavior.

Integrativeness: Students’ wishes to integrate into a L2 community.

Language attitudes: Cognitive, affective, and conative dispositions toward a language.

Ideal L2 self: How students imagine themselves as future language users.

Ought-to L2 selves: How students see what they should accomplish because of outside pressure (parents, teachers, peers, etc.).

Extrinsic and intrinsic motivation: Types of motivation differing in the extent to which the motives are internalised.

Task motivation: Intended effort invested into carrying out a certain activity.

Demotivation: Losing one’s motivation to accomplish something.

Amotivation: Lack of motivation.

Teacher motivation: Subsuming teachers’ own motivation to learn as well as their desire to motivate learners.

The Social-Psychological Phase

Research in the social-psychological phase was guided by Gardner and his colleagues, who developed a complex L2 motivational theory based on social psychological terms, whereby it was assumed that success in language learning depended largely on learners' positive attitudes toward the linguistic cultural community. Gardner (1985, 2006, 2010) presented reviews of several studies conducted in varied contexts that produced evidence that attitudes were indeed key constituents in L2 motivation constructs. The main conceptual result of Gardner's and his associates' efforts was the definition of integrativeness. Gardner (1985, 2001) defined integrativeness in various Canadian contexts as a latent construct made up of the following variables: interest in foreign languages, integrative orientation, and attitudes toward Canadian/European French. As a result of this operationalisation, integrativeness "reflects a genuine interest in learning the second language in order to come closer to the other language community" (Gardner, 2001, p. 5), which can be manifested in either general openness and respect toward the L2 community or actual identification with or integration into the L2 community (Gardner, 2001). Seemingly not much is said about ISLA in Gardner's theory up to this point, but if we take a further step and look at the integrative motive that is composed of attitudinal, goal-directed, and motivational variables, we can see that this motive subsumes integrativeness (as defined in the preceding Key Concepts box), attitudes toward the learning situation (evaluation of the L2 teacher and course) and motivation. Hence, a link is presented between ISLA and motivation by highlighting two important motivational aspects of the classroom: students' attitudes toward the teacher and course. This link is further corroborated in Gardner's socioeducational model on language learning, in which both informal and formal learning contexts were taken into account (Gardner & MacIntyre, 1993), with the latter clearly indicating ISLA environments. According to the socioeducational model, a number of individual difference variables, motivation, and language attitudes included, exert their influence on the linguistic and nonlinguistic outcomes of learning, such as changes in attitudes toward the members of the L2 speech community, in various formal and informal learning contexts. As motivation subsumes attitudinal influences and language attitudes are also integrated into the socioeducational model, the conclusion is that attitudes toward every aspect of ISLA might have a role in shaping students' motivation and hence their ultimate success in language learning (Gardner, 2010).

The Cognitive-Situated Phase

During extended work in Canada, some researchers' interest in the cognitive-situated aspect of L2 motivation set out to broaden the scope of L2 motivation research by incorporating mainstream psychological theories into the field and by offering motivational strategies (i.e., practical implications for L2 teachers for motivating their learners) for classroom teaching. The most important motivational frameworks representing these education-friendly approaches include Crookes and Schmidt's (1991) theory, Dörnyei's (1994) extended motivational framework, and Williams and Burden's (1997) framework. What is common in these three models is that they all contain references to ISLA contexts. Crookes and Schmidt (1991) argued that the classroom level of motivation includes interest, activities, relevance, need for affiliation, feedback, the issue of extrinsic rewards, the effect of students' self-perception, as well as past experiences, as important motivational factors. Dörnyei (1994) highlighted three components related to the *learning-situation level* (educational dimension) of his model that are associated with situation-specific motives rooted in various aspects

of language learning in a classroom setting. Within this level three main types of motivational sources can be separated out: (1) *course-specific motivational components*, which are related to the syllabus, the teaching materials, the teaching method, and the learning tasks; (2) *teacher-specific motivational components*, which are related to the teacher's behaviour, personality, and teaching style; and (3) *group-specific motivational components*, which are related to the dynamics of the learner group. Finally, Williams and Burden's (1997) framework of L2 motivation has been broken down into several factors along the organising principle of external/internal dimensions. The external dimension involved, among others, the role of school environment including comfort, resources, time of day, week, year, size of class and school, and class and school ethos (Williams & Burden, 1997). Despite the fact that these models were informing and inspiring empirical research, very few of them concentrated on the actual classroom level and instead general interest in the field shifted to the temporal dimension of L2 motivation, which took into account the often neglected fact that foreign language learning is a long and arduous enterprise.

The Process-Oriented Phase

From the late 1990s on, researchers have called attention to the changing nature of L2 motivation (Dörnyei & Ottó, 1998; Ushioda, 1998); because learning an L2 is a long enterprise, students' level of motivation is bound to change throughout the process. An example of investigating the temporal dimension of L2 motivation was offered by Ushioda's (1998, 2001) based on a qualitative longitudinal study among university students. According to Ushioda, the process of motivation is basically shaped by either motivation deriving from past experiences (e.g., positive L2 learning or L2-related experiences) or, by motivation directed toward future goals (e.g., personal goals, short-term incentives, language-related goals). Both of these issues can be easily related to ISLA.

Another example of including time as a variable into L2 motivation is the theoretical model proposed by Dörnyei and Ottó (1998). Drawing on Heckhausen and Kuhl's Action Control Theory (e.g., Heckhausen, 1991; Heckhausen & Kuhl, 1985), the motivational process has been broken down into discrete temporal segments by including preactional, actional, and postactional phases and attaching motivational influences and action sequences to each stage. The model describes how initial wishes and desires are first transformed into goals and then into intentions, and how these intentions are acted on, leading to the accomplishment of the goal and/or the termination of action. The process is concluded by the final evaluation. Although not expressed specifically, each stage of the model contains elements highly relevant for L2 motivation in instructed settings (see Table 23.1).

Table 23.1 Classroom-related elements from Dörnyei and Ottó's (1998) process model of L2 motivation

<i>Preactional stage</i>	<i>Actional stage</i>	<i>Postactional stage</i>
<ul style="list-style-type: none"> • Characteristics of the classroom goal structure, both longer and shorter term goals • Attitudes and values in relation to the learning process • Environmental support or hindrance 	<ul style="list-style-type: none"> • Appraisal of the learning process • Teachers' and parents' roles • Reward and goal structure in the classroom • Group dynamical influences 	<ul style="list-style-type: none"> • Characteristics of feedback, praise and received grade

Note: Based on Dörnyei (2001, p. 22).

Current Issues

New impetus was given to L2 motivation research in recent years for at least two reasons. First, Dörnyei's L2 Motivational Self System (Dörnyei, 2005) theory with its parsimonious conceptualisation of L2 motivation has resulted in a large number of studies in different contexts, which were collected in an edited volume by Dörnyei and Ushioda (2009). Second, the emerging development in applied linguistics concerning the theory and application of dynamic system theory (DST) to language learning in general and L2 motivation in particular has also led to a great variety of new research projects compiled in a the recent collection by Dörnyei et al. (2015).

Dörnyei's (2005) L2 Motivational Self System theory posits that students' motivated learning behavior (i.e., how much effort they are willing to invest into language learning and how persistent they are) will be largely affected by three distinct variables: their ideal L2 self, that is, to what extent students can imagine themselves as highly proficient users of the given foreign language; their ought-to L2 self, which describes what outside pressures students acknowledge throughout the learning process; and finally, their language learning experience, which influences attitudes toward the classroom processes (Dörnyei, 2005; Dörnyei & Ushioda, 2009). A number of studies justified the validity of this tripartite theory (Dörnyei & Ushioda, 2009) but the three parts have received various amounts of emphasis in subsequent research.

The central tenet of this theory has become students' ideal L2 selves, which led research into vision-related issues. Dörnyei and Kubanyiova (2014) devoted a volume to the use of vision in the classroom. The main point of their book is that by adding a vision as a part of teaching, students are better able to develop future self-guides, such as students' ideal L2 selves, and goals that will enhance their motivation and as a result, their achievement in language learning. Dörnyei and Kubanyiova (2014) also argue that their teachers' visions about themselves as language learners (i.e., when teachers are (or used to be) L2 learners of the students' target language themselves) as well as language teachers contribute to the motivational dynamics of the L2 classroom. The role of ought-to L2 self has been seen as very complex because the diverse nature of outside expectations makes the operationalisation of the concept rather difficult (Kormos & Csizér, 2008). As for the third component of the model, language learning experiences (unspecified whether or not these experiences are related to instructed language learning) remain a somewhat neglected aspect of the model, with some studies measuring it as general positive attitudes to learning while others as positive attitudes to classroom learning (You, Dörnyei, & Csizér, 2016).

The most recent development in L2 motivation research is the inclusion of dynamic system theory dynamic system theory into the L2 motivation field. Dörnyei et al.'s (2015) volume on the applicability of and empirical evidence on dynamic system theory points toward the fact that a classroom can indeed be selected as a "domain of reality" (p. 424) to investigate L2 motivation because L2 motivation in the classroom "is nothing if not complex and dynamic" (p. 421). The volume indeed contains several classroom-related motivational studies. Waninge (2015) looks into motivational and demotivational attractor states in classrooms, that is, states when motivation stops fluctuating for a time and how these relatively stable states are linked to language learning experience. This study concludes that there are four main characterising elements related to classroom experience: interest, boredom, neutral attention, and anxiety. Changes in motivation during a 14-week writing seminar are reported (Piniel & Csizér, 2015), in which it was mapped how the motivation, anxiety, and self-efficacy changed during a university writing course. Learners with different level of

motivation are typified and investigated (Chan, Dörnyei, & Henry, 2015) and it was pointed out that a limited number of learners' archetype existed in a class, which teachers were usually aware of. MacIntyre and Serroul's (2015) investigation delves into L2 motivation within a short period of time and looked at how it fluctuated second by second during a particular task. The conclusion of Dörnyei et al.'s (2015) volume is that because theoretical considerations of dynamic system theory take into account both change and interaction of various motivational variables in the classroom, dynamic system theory researchers could offer important contributions to L2 motivation research in ISLA settings.

Empirical Evidence

There are four areas of L2 motivation research that have provided empirical evidence related to ISLA and classroom-related processes. In the following I give a brief summary of the factors affecting learner motivation, including: teachers, task, peers, and demotivation. Although the effect of curriculum and teaching methods can also be considered as part of the classroom context (Dörnyei, 2009), there are hardly any studies on them, and therefore they should be considered issues for future research.

The Role of Teachers in Students' Motivation

Concerning the role of teachers in shaping students' level of motivated behavior, there are a number of publications on ways teachers might be able to motivate their students. In addition, there are lists of motivational strategies in Brophy (1987), Dörnyei (2001), Dörnyei and Csizér (1998) as well as Cheng and Dörnyei (2007) with the latter two relying on teachers' self-report on the relative importance of the various strategies (for details see the section on Pedagogical Implications). Still, there are markedly fewer empirical studies on how teachers' motivation actually affects students' motivation. In order to fill this research niche, there could be different quantitative and/or qualitative approaches to investigate teachers' impact on students' motivation, but all of these possible studies rely on relatively complex research designs, as both teachers' and students' motivation need to be measured and then matched during data analysis, which might partly explain the scarcity of this type of research.

The classic approach to investigating the influence of teachers on students' motivation is to observe what motivational strategies teachers use and to measure students' motivation simultaneously. These strategies include practical techniques that teachers use to motivate their students, for example, promoting motivational values, cooperation, autonomy, piquing students' curiosity, and effective feedback. In order to do this, Guilloteaux and Dörnyei (2008) developed an instrument called Motivation Orientation in Language Teaching (MOLT), which allows researchers to collect classroom data on teachers' use of motivational strategies as well as students' motivated learning behavior. Their results from South Korea indicated strong positive associations between teachers' motivational strategy use and students' behavior, concluding that "the teachers' motivational practice does matter" (p. 72). MOLT was also used in an Iranian context with similar results (Papi & Abdollahzadeh, 2012), further corroborating the usefulness of MOLT and the importance of L2 motivational strategies in ISLA. More complex results were obtained by Mezei (2014) in a Hungarian context, where it was found that teachers' use of motivational strategies impacted students' motivated learning behavior both directly and indirectly through other important predicting variables such as ideal L2 self, instrumental orientation (pragmatic gains from knowing a

foreign language) and various self-regulatory processes, that is, in what ways learners are able and willing to take responsibility for their own learning (Mezei, 2014).

The effects of possible mediating variables on the impact of teachers' motivation on students' learning have also been investigated in various studies. Bernaus and Gardner (2008) studied the link between teachers' motivational strategy use and students' perception of these strategies and other self-related variables. Their results indicated a mismatch between teacher and student data on reported strategy use, but students' perception of strategy use was positively linked to students' motivation and achievement, indicating the importance of students' perception of teacher behavior in the motivational process. Both Ruesch, Bown, and Dewey's (2012) investigation and Wong's (2014) study added that not only should students' perception of teacher behavior be taken into account, but in their comparison of data from different countries the researchers came to the conclusion that cross-cultural differences and sociocultural milieu also impacted how teachers' behavior affected students' learning. In addition, students' level of L2 knowledge, and their initial motivation also seem to be related to the impact of teachers' motivational strategies on students' motivation (Sugita & Takeuchi, 2010). In a similar vein, Sugita, McEown and Takeuchi (2014) provided further empirical evidence that motivation strategies had differing impact on students with low and high motivation. All these findings, thus, shed light on the important and complex ways in which teachers' conscious motivational practices can play a role in L2 motivation and thus, ultimately, in the learning process.

Another way to investigate teachers' role in students' motivation is a novel line of study linked to Dörnyei's L2 Motivational Self System theory. In these studies researchers designed various intervention programs embedded in regular teaching practices that aimed to enhance students' visions about themselves as future language users. In these intervention programs teachers/researchers used various strategies to help students develop, enhance, and strengthen their ideal L2 selves. Magid (2014) used scripted imagery, that is, students had to imagine their desired and feared future selves based on some guidelines the teacher provided; Letty (2014) employed imagery training strategies; and Mackay (2014) implemented a motivational training program based on Hadfield and Dörnyei (2013). They all found that these intervention programs designed to enhance students' ideal L2 selves and visions about themselves carried positive values for students, and their level of motivation indeed increased.

Apart from using motivational strategies, there are other ways in which teachers might influence students' L2 motivation. Noels, Clément, and Pelletier (1999), for example, investigated how teachers' communicative style affected L2 motivation. Their results implied that the extent to which students internalise various motives, that is their extrinsic and intrinsic motivation, was differently affected by the teacher's communicative style. Intrinsic motivation correlated negatively with a controlling communicative style but correlated positively with an informative communicative style; extrinsic motivation, on the other hand, did not seem to be affected by teachers' communicative style. As a consequence, Noels et al. (1999) have reached the conclusion that "by interacting with students in ways that develop their autonomy and competence, teachers may change the students' type of motivation, and thereby contribute to better learning" (p. 31).

Task Motivation

A highly promising classroom-based research direction involves the investigation of task motivation, which was considered to be the most "situation-specific" paradigm possible in the L2 motivation field (Dörnyei, 1996; Julkunen, 2001). Task motivation explains why

students behave as they do in a specific learning situation where they are carrying out a specific task (Dörnyei, 2002). Accordingly, Julkunen (2001) defines classroom motivation “as a continuous interaction process between the learner and the environment” (p. 29). Despite possible relevance to classroom motivation, there are only a handful studies on task motivation. Julkunen (1989) proved that a co-operative task environment, as opposed to individual or competitive situations, was the most motivating for both high- and low-achievers. Dörnyei and Kormos (2000) as well as Dörnyei (2002) found that aspects of L2 motivation affected the execution of a task and concluded that task motivation was co-constructed by task participants. Dörnyei and Tseng (2009) investigated how students’ motivation affected task engagement concerning vocabulary learning in light of the experience of the teacher by including into the research design the comparison of novice and experienced teachers’ task-related practices in class. They proposed and empirically tested a tripartite system that represented task motivation by including “task execution, task appraisal, and action control, which result in students’ engagement in the task, their evaluation of the process of task completion as well as self-regulating task completion” (Dörnyei & Tseng, 2009, p. 119). Their results, based on structural equation modelling, validated a circular relationship among the three constructs with novice and experienced learners behaving slightly differently, indicating that novice teachers had problems with monitoring students while they were completing the various tasks. Dörnyei and Tseng propose that “the quality of motivational task processing is indicative of the quality of the SLA process” (2009, p. 122), and therefore suggest that future research should also concentrate on how motivational task-processing relates “to attention, noticing as well as implicit/explicit or incidental/intentional learning” (p. 122). In a similar vein, Csizér and Tankó (in press) investigated the relationship between an academic writing task and L2 motivation. Their cross-sectional investigation indicated a positive link between successful task completion and students’ reported level of L2 motivation. In addition, it was also found that other individual variables, such as anxiety and self-regulation, also contributed to successful task completion (Tankó & Csizér, 2014). Furthermore, based on these results it can also be concluded that there is a close link between students’ regulating task completion and their level of motivation, with more motivated students being more willing to take responsibility for the learning process in general and the task at hand in particular (Csizér & Tankó, in press).

Another example of task-based motivation is offered by MacIntyre and Serroul (2015), who positioned their study in the dynamic system theory and investigated task motivation on a “per-second timescale” (p. 109). The study, which involved L2 learners completing eight different speaking tasks, showed that their motivation indeed changed throughout task completion. Both approach and avoidance motivation were described, that is, whether or not the participants were willing to complete the task or wanted to avoid it. The level of motivation was based on students’ perception of task difficulty, necessary vocabulary for successful task completion as well as grammar-related issues. In addition, significant positive correlation was found between students’ initial assessment of their own task motivation and their actual motivation while performing the tasks.

Group Dynamics

As instructed language learning often happens in groups, it is logical to assume that group-related variables might affect students’ motivation and thus learning behavior and achievement. Unfortunately, despite the fact that group dynamics is an established field in social psychology and there are theoretical contributions to the L2 motivation field (Dörnyei &

Murphey, 2003), the number of empirical studies is low (although group-related L2 motivation strategies are included among L2 motivation strategies). Clément, Dörnyei, and Noels (1994) found that perceived cohesiveness of the group affected the motivational construct and correlated with students' intended effort. Ghaith (2003) investigated how different types of learning modes (cooperative, individual, and competitive) shaped classroom climate. Although this study did not address L2 motivation in a direct way, a significant relationship was found between cooperation ("learners work together in small groups to achieve common goals"; Ghaith, 2003, p. 84) and group cohesiveness ("students enjoy working with their classmates because they know them and consider them friends"; Ghaith, 2003, p. 85), with the latter related to L2 motivation. Chang (2010), who investigated how group-related variables contribute to students' motivation, found that correlational evidence existed between group cohesiveness and group norms on the one hand, and language learning motivational processes on the other. Qualitative data corroborated these results indicating the students were aware of how classmates could motivate or, in a more unfortunate situation, demotivate one another.

Demotivation

Another potentially pivotal issue in ISLA concerns the empirical investigation of students' demotivation, that is, students losing their motivation during the learning process. Unlike amotivation in self-determination theory (e.g., Noels, 2001), which implies complete lack of motivation, demotivation is a process whereby initially motivated students lose their willingness to invest energy into language learning. Early research showed that both internal and external factors could contribute to demotivation. Chambers (1993) found that demotivated students typically lacked self-confidence, did not see the importance of language learning and had conflicts with their teachers. Oxford (1998), based on her qualitative investigation, identified two main demotivating issues, namely teaching methods and learning tasks—both clearly associated with ISLA. Similarly, Ushioda (1998) concluded that demotivation is linked to ISLA, namely to teaching methods and learning tasks. In Dörnyei's (1998) study teachers were identified to be the most "important" demotivating factors in students' motivation.

After the initial interest in demotivation, several studies emerged from different contexts using various data collection methods to investigate demotivation. In Hungary, Nikolov (2001) found that classroom-related processes, more specifically teachers, played an important role in shaping students' dispositions, motivation, and achievement. In the Japanese contexts demotivating instances were found to be linked to classroom-related issues, such as teachers, classroom characteristics, and classroom environment (Sakai & Kikuchi, 2009). Based on another Japanese study, Falout, Elwood, and Hood (2009), pointed out that there was a relationship between students' level of proficiency and some characteristic demotivating factors, that is, less proficient learners were less able to cope with demotivating instances in the classrooms. In Vietnam, Trang and Baldauf (2007) used a stimulated recall essay task to investigate demotivation and found that teachers contributed to students' demotivation by selecting teaching methods not suited to learners' learning styles, for example visual style.

Apart from describing demotivational instances, researchers found two other important issues to consider. First, Falout et al. (2009) investigated reactive factors, that is, how students reacted to demotivating instances. Second, Kim (2011) underlined the importance of perception, that is, not the context itself that was defining but how students' viewed external demotivating instances, that is how and why students recognised these instances and dealt

with them. These results indicate the active role students need to play in their learning process whereby they try to take responsibility for their own motivation. Still, it seems that in terms of ISLA the pivotal role of teachers needs to be underlined and pedagogical implications considered in motivating students, as both the selected teaching methods and teacher behavior can be demotivating to learners.

Pedagogical Implications

Pedagogical implications based on empirical results are difficult to form for several reasons. First, researchers are often interested in general motivation characteristics that do not translate well into the practicalities of various instructional contexts. Second, as Dörnyei and Ushioda (2011) point out, the L2 motivation research might lack “a level of sophistication that would allow scholars to translate research results into straightforward educational recommendations” (p. 104), but even if it were possible to offer such recommendations they would need to be adapted to educational micro-contexts, that is, different classrooms (Holliday, 1994). Nonetheless, based on theoretical considerations, one can propose a number of pedagogical implications.

As part of the process model of L2 motivation, Dörnyei (2001) offered a conceptualisation of the motivational teaching practice, which includes four main components: creating the basic motivational conditions, generating initial motivation, maintaining and protecting motivation, and encouraging positive retrospective self-evaluation. This framework is complemented by 35 motivational strategies for classroom use (Dörnyei, 2001), ranging from setting a personal example of enthusiasm to creating an ideal context for the learning. The list is admittedly daunting; therefore, Dörnyei urges teachers to take a stepwise approach when incorporating the strategies into their teaching practice.

Another line of research with fruitful pedagogical implications grew out of the vision-related theoretical work. Dörnyei and Kubanyiova (2014) propose that the most important pedagogical intervention could be vision-related motivational impact on students: “we have come to believe that vision is one of the single most important factors within the domain of language learning: where there is a vision, there is a way” (p. 2). They dedicated a full volume to exploring vision-related pedagogical implications both for students and teachers alike because “we understand vision to be one of the highest-order motivational forces, one that is particularly fitting to explain the long-term, and often life-long, process of mastering a second language” (p. 4). Describing possible ways to motivate students takes up six chapters in the book, which deal with how vision can help students at various stages of learning and what teachers might be able to do to develop and maintain students’ visions (Dörnyei & Kubanyiova, 2014). The first step should be helping students create visions for themselves by providing guided imagery and narratives. Next, the vision needs to be developed and strengthened with the help of vision inducing tasks, such as learning journals, virtual tools, and strengthening group vision. Third, the created vision needs to be rendered realistic for the learners in order for them to strive to reach their vision. Fourth, it is important to transform vision into action in order to enhance L2 motivation and ultimate achievement. Fifth, the vision needs to be maintained by helping students with reminders and possible adjustments to their visions. Finally, failure needs to be considered not only by drawing students’ attention to possible negative outcomes but also helping students develop realistic external motivational drives to succeed.

Based on the overviewed empirical evidence, it can be concluded that there is no single tip that would work for each teacher in every context. It does not mean that teachers should

not be aware of their roles as motivators, as I argued in this chapter that motivation is a key ingredient to successful classroom learning. Based on the empirical evidence, there are five major issues that seem to be highly relevant to ISLA.

Teaching Tips

- Set an example as being motivated by getting to know your students' interest and how it can be incorporated into ISLA.
- Be aware of the group dynamic processes: how they can help/hinder the learning processes.
- Know that task motivation is an important part of motivation: even during a relatively short task students' motivation can ebb and flow.
- Do not be afraid of demotivation: it will happen in the classroom. Try raising students' awareness and show how they can turn demotivation around and motivate themselves.
- Create visions of the goals you would like to achieve with your students.

Future Directions

Based on the overview presented in this chapter, it is clear that despite the fact that the L2 motivation research field has been fast developing in multiple directions, there are a number of classroom-related issues that need further investigation. Before going into detail, I have to point out that both large-scale quantitative studies and longitudinal qualitative studies could contribute to a better understanding of how instruction can impact learners' motivation. In order to further our understanding of ISLA, more research should be carried out on the relationship of classroom-related variables and L2 motivation. As perception itself seems to be an important issue in instructed learning, self-reported data accompanied by observation should not be ignored. In addition, Dörnyei and Ushioda (2011) propose a dynamic investigation of students' motivation and the daily events of a language course.

I think there is an increased need for research into the effects of instruction on L2 to find out how language instruction in general, and specific aspects of language instruction in particular, impact L2 learners. In addition, the actual usefulness of L2 motivation strategies in motivating language learners in various classrooms could be further investigated by linking various students' characteristics to the efficiency of strategies. A possible research direction could be to conduct case studies that would involve classroom observation with the aim of looking at how teachers use motivational strategies, and compare their practice to motivational data from students. The complex link needs to be further explored by adding possible mediating variables into the picture that impact the link between teachers' classroom behavior and students' perception. In addition, teachers' teaching styles in general, and communication styles in particular, can add further information to the growing body of evidence on how teachers' behavior might impact students' motivation (Dörnyei & Ushioda, 2011).

More research is needed to investigate not only students' general motivational dispositions but also task-specific motivation: what are their views on various learning tasks and how do task-related characteristics shape their motivation. Task motivation could be investigated in longitudinal studies, as was suggested by Dörnyei and Ushioda (2011), and the dynamic system theory paradigm could also be taken into account (Dörnyei et al., 2015).

Group dynamics, the role of teachers and various methods of instruction should be taken into account as well. Quantitative studies concentrating on different classroom-related variables can help us to understand how the development and level of L2 motivation changes. Qualitative studies are well suited to map how group and interpersonal relationships might shape students' motivation and achievement (Dörnyei & Ushioda, 2011). Longitudinal case studies can contribute to our understanding of how the development of groups can enhance and/or hinder L2 motivation. Within the field of group dynamics, goal-related issues can also be investigated: how goals of individual students can contribute to group goals and how common goals within a group can help L2 motivation.

As outlined earlier, demotivation is a complex issue and more research could shed light on three important domains: (1) demotivation and its relationship to general motivational dispositions and personal characteristics; (2) demotivation and its situation specificity; and (3) issues related to the valid measurement of motivation in the face of possible demotivation (Dörnyei & Ushioda, 2011). The fact that most results predicted teachers to be the most important demotivating factor in students' motivation further highlights the necessity of investigating the relationship between the ways teachers motivate their students and its impact on student motivation.

There is a need for more longitudinal and ethnographic studies to investigate the intricate relationship between L2 instruction, L2 motivation, and L2 learning. In addition, qualitative studies could look into differences in L2 motivational processes in instructed and noninstructed learning contexts, which could be potentially important for uncovering differences concerning English as a global language and other regionally important languages. In addition, as pointed out earlier, the effect of curriculum and teaching method on shaping L2 motivation could also be investigated.

Within the dynamic system theory paradigm, many of the situation-specific issues could be researched (see Dörnyei et al., 2015). In addition, instructed language teaching research could benefit from more specific goal-related studies. Despite the fact that goals are a thoroughly researched field in psychology, we still know relatively little about how short- and long-term goals might affect the L2 learning process in instructed contexts, including individual, classroom, and school-level goals. In addition, as dynamic system research takes time as a variable into consideration, motivational change in classrooms could be further explored: how L2 motivation changes during a task, a lesson, a week, a month and school year, and even in longer periods could enhance our knowledge on important issues shaping L2 motivation.

Conclusion

Based on this brief overview of classroom motivation, it can be concluded that despite the fact that L2 motivation is a much-researched field, there is still a lot to do in the investigation of ISLA and L2 motivation. Taking differing contexts and time as factors into consideration, I am sure that L2 motivation research will stay in mainstream research in both applied linguistics and language pedagogy. Sensitivity to issues related to L2 motivation and achievement will be able to further fine-tune research studies in order to inform researchers and teachers alike.

References

- Bernaus, M., & Gardner, R. C. (2008). Teacher motivation strategies, student perceptions, student motivation and English achievement. *Modern Language Journal*, 92, 387–401.
- Brophy, J. E. (1987). Synthesis of research on strategies for motivating students to learn. *Educational Leadership*, 45, 40–48.

- Candlin, C. N., & Mercer, N. (Eds.). (2001). *English language teaching in its social context*. London: Routledge.
- Chambers, G. N. (1993). Taking the 'de' out of demotivation. *Language Learning Journal*, 7, 13–16.
- Chan, L., Dörnyei, Z., & Henry, A. (2015). Learner archetypes and signature dynamics in the language classroom: A retrodictive qualitative modelling approach to studying L2 motivation. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 238–259). Bristol: Multilingual Matters.
- Chang, L. (2010). Group processes and EFL learners' motivation: A study of group dynamics in EFL classrooms. *TESOL Quarterly*, 44, 129–154.
- Cheng, H.-F., & Dörnyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching*, 1(1), 153–174.
- Clément, R., Dörnyei, Z., & Noels, K. (1994). Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language Learning*, 44, 417–448.
- Crookes, G., & Schmidt, R. (1991). Motivation: Reopening the research agenda. *Language Learning*, 41, 469–512.
- Csizér, Kata, & Magid, M. (2014). *The impact of self-concept on language learning*. Bristol: Multilingual Matters.
- Csizér, Kata., & Tankó, G. (in press). English majors' self-regulatory control strategy use in academic writing, and its relation to L2 motivation. *Applied Linguistics*.
- Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *Modern Language Journal*, 78, 273–284.
- Dörnyei, Z. (1996). Moving language learning motivation to a larger platform for theory and practice. In R. Oxford (Ed.), *Language learning motivation: Pathways to the new century* (Technical Report #11, pp. 71–80). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center.
- Dörnyei, Z. (1998, March). *Demotivation in foreign language learning*. Paper presented at the TESOL '98 Congress, Seattle, WA.
- Dörnyei, Z. (2001). *Motivational strategies*. Cambridge: Cambridge University Press.
- Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (ed.), *Individual differences and instructed language learning* (pp. 137–158). Amsterdam, the Netherlands: John Benjamins.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning*, 59, 230–248.
- Dörnyei, Z., & Csizér, K. (1998). Ten commandments for motivating language learners: Results of an empirical study. *Language Teaching Research*, 2, 203–229.
- Dörnyei, Z., Csizér, K., & Németh, N. (2006). *Motivational dynamics, language attitudes and language globalisation: A Hungarian perspective*. Clevedon: Multilingual Matters.
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4, 275–300.
- Dörnyei, Z., & Kubanyiova, M. (2014). *Motivating learners, motivating teachers: Building vision in the language classroom*. Cambridge: Cambridge University Press.
- Dörnyei, Z., MacIntyre, P., & Henry, A. (Eds.). (2015). *Motivational dynamics in language learning*. Bristol: Multilingual Matters.
- Dörnyei, Z., & Murphey, T. (2003). *Group dynamics in the language classroom*. Cambridge: Cambridge University Press.
- Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics* (Thames Valley University, London), 47, 173–210.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York: Routledge.
- Dörnyei, Z., & Tseng, W.-T. (2009). Motivational processing in interactional tasks. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp. 117–134). London: Routledge.

- Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Harlow: Longman.
- Falout, J., Elwood, J., & Hood, M. (2009). Demotivation: Affective states and learning outcomes. *System*, 37, 403–417.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
- Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (Technical Report #23, pp. 1–19). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Gardner, R. C. (2006). The socio-educational model of second language acquisition: A research paradigm. In S. H. Foster-Cohen, M. Medved Krajnovic, & J. Mihaljevic Djigunovic (Eds.), *Eurosla Yearbook Volume 6* (pp. 237–260). Amsterdam, The Netherlands: John Benjamins.
- Gardner, R. C. (2010). *Motivation and second language acquisition: The socio-educational model*. New York: Peter Lang.
- Gardner, R. C., & MacIntyre, P. D. (1993). A student's contribution to second-language learning. Part II: Affective variables. *Language Teaching*, 26, 1–11.
- Ghaith, G. (2003). The relationship between forms of instruction, achievement and perceptions of classroom climate. *Educational Research*, 45, 89–93.
- Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 42, 55–77.
- Hadfield, J., & Dörnyei, Z. (2013). *Motivating learning*. Harlow: Longman.
- Heckhausen, H. (1991). *Motivation and action*. New York: Springer.
- Heckhausen, H., & Kuhl, J. (1985). From wishes to action: Dead ends and short cuts on the long way to action. In M. Frese & J. Sabini (Eds.), *Goal-directed behaviour: The concept of action in psychology* (pp. 134–160). Hillsdale, NJ: Lawrence Erlbaum.
- Holliday, A. R. (1994). *Appropriate methodology and social context*. Cambridge: Cambridge University Press.
- Julkunen, K. (1989). *Situation- and task-specific motivation in foreign-language learning and teaching*. Joensuu: University of Joensuu.
- Julkunen, K. (2001). Situation- and task-specific motivation in foreign language learning. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (Technical Report #23, pp. 29–41). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Kim, T.-Y. (2011). Sociocultural dynamics of L2 learning (de)motivation: An activity theory analysis of two adult Korean immigrants. *Canadian Modern Language Review*, 67, 91–122.
- Kormos, J., & Csizér, K. (2008). Age-related differences in the motivation of learning English as a foreign language: Attitudes, selves and motivated learning behaviour. *Language Learning*, 58, 327–355.
- Letty, C. (2014). Effects of an imagery training strategy on Chinese university students' possible second language selves and learning experiences. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 357–376). Bristol: Multilingual Matters.
- MacIntyre, P., & Serroul, A. (2015). Motivation on a per-second timescale: Examining approach-avoidance motivation during L2 task performance. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 109–138). Bristol: Multilingual Matters.
- Mackay, J. (2014). Applications and implications of the L2 motivational self system in a Catalan EFL context. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 377–400). Bristol: Multilingual Matters.
- Magid, M. (2014). A motivational programme for learners of English: An application of the L2 motivational self system. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 333–356). Bristol: Multilingual Matters.

- Mezei, G. (2014). The effects of motivational strategies on self-related aspects of student motivation and second language learning. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 289–309). Bristol: Multilingual Matters.
- Nikolov, M. (2001). A study of unsuccessful language learners. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 149–170). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Noels, K. (2001). New orientations in language learning motivation: Towards a model of intrinsic extrinsic, and integrative orientations and motivations. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (Technical Report #23, pp. 43–68). Honolulu: University of Hawai'i, Second Language & Curriculum Center.
- Noels, K., Clément, R., & Pelletier, L. G. (1999). Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation. *Modern Language Journal*, 83, 23–34.
- Oxford, R. (1998, March). *The unravelling tapestry: Teacher and course characteristics associated with demotivation in the language classroom: Demotivation in foreign language learning*. Paper presented at the TESOL '98 Congress, Seattle, WA.
- Papi, M., & Abdollahzadeh, E. (2012). Teacher motivational practice, student motivation, and possible L2 selves: An examination in the Iranian EFL context. *Language Learning*, 62, 571–594.
- Piniel, K., & Csizér, K. (2015). Changes in motivation, anxiety, and self-efficacy during the course of an academic writing seminar. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 164–194). Bristol: Multilingual Matters.
- Ruesch, A., Bown, J., & Dewey, D. P. (2012). Student and teacher perceptions of motivational strategies in the foreign language classroom. *Innovation in Language Learning and Teaching*, 6(1), 15–27.
- Sakai, H., & Kikuchi, K. (2009). An analysis of demotivators in the EFL classroom. *System*, 37, 57–69.
- Sugita McEown, M., & Takeuchi, O. (2010). What can teachers do to motivate their students? A classroom research on motivational strategy use in the Japanese EFL context. *Innovation in Language Learning and Teaching*, 4(1), 21–35.
- Sugita McEown, M., & Takeuchi, O. (2014). Motivational strategies in EFL classrooms: How do teachers impact students' motivation? *Innovation in Language Learning and Teaching*, 8(1), 20–38.
- Tankó, G., & Csizér, K. (2014). Investigating English majors' individual differences through their argumentative essays. In V. Ruttkay & B. Gárdos (Eds.), *HUSSE II proceedings of the 11th conference of the Hungarian society for the study of English* (pp. 701–718). Budapest: L'Harmattan.
- Trang, T. T. T., & Baldauf, R. B. (2007). Demotivation: Understanding resistance to English language learning—The case of Vietnamese students. *Journal of Asia TEFL*, 4(1), 79–105.
- Ushioda, E. (1998). Effective motivational thinking: A cognitive theoretical approach to the study of language learning motivation. In E. A. Soler & V. C. Espurz (Eds.), *Current issues in English language methodology* (pp. 77–89). Castelló de la Plana, Spain: Universitat Jaume I.
- Ushioda, E. (2001). Language learning at university: Exploring the role of motivational thinking. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (Technical Report # 23, pp. 93–125). Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center.
- Waninge, F. (2015). Motivation, emotion and cognition: Attractor states in the classroom. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 195–213). Bristol: Multilingual Matters.
- Williams, M., & Burden, L. R. (1997). *Psychology for language teachers*. Cambridge: Cambridge University Press.
- Wong, R. M. H. (2014). An investigation of strategies for student motivation in the Chinese EFL context. *Innovation in Language Learning and Teaching*, 8(2), 132–154.
- You, C. J., Dörnyei, Z., & Csizér, K. (2016). Motivation, vision, and gender: A survey of learners of English in China. *Language Learning*, 66(1), 94–123.

Psychological Dimensions and Foreign Language Anxiety

Jean-Marc Dewaele

Background

Few psychological dimensions have been as intensively researched in SLA as anxiety. As Dörnyei and Ryan (2015) put it, anxiety has been in the limelight of SLA research for several decades. Indeed, learners, teachers, and researchers agree that anxiety is a common experience and they have been interested in knowing to what extent anxiety inhibits language learning and language production. This question fitted squarely in the more general research into the internal characteristics of the “good language learner” in the mid-1970s. Naiman, Fröhlich, Stern, and Todesco (1978) looked at 72 Anglo-Canadian high school students learning French as a second language (L2) who scored highest on the Listening Test of French Achievement and an Imitation Test and tried to determine whether these “good language learners” had a unique psychological profile, similar motivations, attitudes, cognitive styles, or learning strategies. It turned out that good language learners, like self-made millionaires, have positive attitudes and strong motivation but differ widely in personality profiles. The latter was so unexpected that the authors concluded—rather surprisingly—that the lack of correlations between the dependent variables and personality traits was due to the instruments for measuring personality and cognitive traits lacking construct validity (see Dewaele & Furnham, 1999 for a closer analysis). Naiman et al. (1978) never wondered whether their own research design was to blame for the lack of significant relationships, especially their choice of L2 measures based on written performance. Interestingly, the (lack of) anxiety did not appear as a distinctive characteristic of good language learners. Based on the feedback received from participants to open questions about their learning behaviour and personality, Naiman et al. (1978) concluded that good language learners were meticulous, sociable, independent, and persevering—but not anxiety-free.

One of the difficulties of presenting existing research on psychological dimensions and Foreign Language Anxiety (FLA) is that all variables had been operationalized and measured in different ways, which led to confusing results when research started in the 1970s (MacIntyre, in press). I will show how SLA researchers adopted a broad framework in the 1980s that has been used and refined ever since. Personality psychologists have also operationalized and measured a plethora of personality traits, states, and facets of personality

traits using a wide range of approaches and instruments, which falls outside the scope of the present chapter. I will therefore refer to personality constructs that seem widely accepted in the field, and will pay particular attention to psychological dimensions that have been linked to FLA. In reviewing the literature, I will follow Plonsky and Oswald's (2014) recent reinterpretation of effect sizes in SLA research.¹

The Confounded Approach in Foreign Language Anxiety (FLA) Research

The first studies into the effects of anxiety on SLA (Chastain, 1975; Kleinmann, 1977; Swain & Burnaby, 1976; Tucker, Hamayan, & Genesee, 1976) gave contradictory results. In his early review of the literature Scovel (1978) observed:

The research into the relationship of anxiety to foreign language learning has provided mixed and confusing results, immediately suggesting that anxiety itself is neither a simple nor well-understood psychological construct and that it is perhaps premature to attempt to relate it to the global and comprehensive task of language acquisition.

p. 132

In his recent overview of language anxiety research and trends, MacIntyre (in press) described this first phase of research as the Confounded Approach “because the ideas about anxiety and their effect on language learning were adopted from a mixture of various sources without detailed consideration of the meaning of the anxiety concept for language learners” (n.p.). The heart of the problem was, according to MacIntyre, the fact that “not all types of anxiety that can be defined and measured are likely to be related to language learning” (n.p.). Scovel (1978) tried to explain the inconsistent results by distinguishing, on the one hand, facilitating and debilitating anxiety, and, on the other hand, trait and state conceptualizations of anxiety, namely the general tendency to experience anxiety across situations (trait) and the more occasional experience of feeling anxious in specific situations (state) (cf. Spielberger, 1966). MacIntyre (in press) argued that the distinction between facilitating and debilitating anxiety has “not been a particularly useful path for SLA research, but the trait/state distinction has been conceptually solid” (n.p.).²

The Specialized Approach in Foreign Language (Classroom) Anxiety Research

The second phase of anxiety research in SLA, according to MacIntyre (in press), was the Specialized Approach, which started with the publication of Horwitz (1986) and Horwitz, Horwitz, and Cope (1986). The authors were influenced by Gardner's suggestion (1985, p. 34) that “the conclusion seems warranted that a construct of anxiety which is not general but instead is specific to the language acquisition context is related to second language achievement.” Gardner argued for a reorientation of the conceptualization and measurement of anxiety in SLA and contributed himself to this second phase of research in the late 1980s in collaboration with MacIntyre.

Horwitz et al. (1986) developed the construct of (Foreign/Second) Language Anxiety that reflected an individual's tendency to be anxious in the specific situation of language learning. Horwitz (in press) explained that

specific anxieties have characteristics of both trait and state anxieties. When individuals experience Language Anxiety, they have the trait of feeling state anxiety when

participating in language learning and/or use. It is also likely that individuals who experience Language Anxiety would feel anxious simply thinking about language learning and/or use.

n.p.

Horwitz et al. (1986) included descriptions of three specific anxieties: Communication Apprehension (anxiety about (public) speaking), Test Anxiety (anxiety experienced in testing situations or in anticipation of testing situations), and Fear of Negative Evaluation (the fear that people will judge the learner negatively) to illustrate concepts of specific anxieties. Horwitz (in press) explained that these three related anxieties were merely examples of specific anxieties, not the three unique components of foreign language classroom anxiety, as it was assumed in later research (cf. Aida, 1994).

Horwitz (1986) developed the argument that Language Anxiety was only analogous to—and not composed of—the three related anxieties. She described the development and validation of the 33-item Foreign Language Classroom Anxiety Scale (FLCAS). The items came from a number of sources including the experiences of anxious language learners. Internal consistency for the FLCAS, measured by Cronbach's alpha, was high (.93). In order to demonstrate the independence of FLCA from previously reported specific anxieties, Horwitz calculated the correlations between her FLCA scores and other types of anxieties such as Trait Anxiety, Communication Apprehension, Test Anxiety, and Fear of Negative Evaluation.³ Horwitz's aim was to demonstrate how small the overlap was between FLCA and the three analogous anxieties “in order to establish the construct validity of a scale designed to elicit foreign language anxiety” (Horwitz, in press, n.p.). She found a nonsignificant correlation of $r = .28$ ($p = .063$) between the FLCAS and the Personal Report of Communication Apprehension, and significant correlations of $r = .36$ ($p < .007$) between the FLCAS and the Fear of Negative Evaluation Scale and the Test Anxiety Scale ($r = .53$, $p < .001$). These results suggest a moderate effect, with 13% of explained variance for the first correlation analysis and 28% for the second analysis. Horwitz (1986) argued that the results supported the contention that FLA could be discriminated from the related constructs but admitted that a moderate association existed with test anxiety. She also found a significant positive correlation of the FLCAS with the Trait scale of the State-Trait Anxiety Inventory (Spielberger, 1983) ($r = .29$, $p < .002$), which represents a small effect size with 8.4% of explained variance. Looking back at her original study, Horwitz concluded that “people who are generally anxious in their lives may be slightly more likely to be anxious in language learning. This finding also means that some anxious language learners do not experience a general tendency to anxiety in their daily lives” (Horwitz, in press, n.p.). She concluded that the amounts of shared variance between the FLCAS and the other anxiety measures were small enough to support “the construct validity of the FLCAS and the existence of Language Anxiety as a specific anxiety independent of other types of anxiety” (n.p.).

MacIntyre and Gardner (1989) collected data from 104 Anglo-Canadian students who had French as an L2 and used factor analysis on various anxiety scales (Trait Anxiety Scale, State Anxiety, Test Anxiety, Computer Anxiety Scale, specific Classroom Anxieties [measuring anxiety in classes of Mathematics, French L2 and English L1], French Use Anxiety Scale, and Audience Sensitivity). The factor analysis yielded a two-factor solution that accounted for 48% of the variance. Factor 1 was labelled General Anxiety after showing high loadings from the Trait Anxiety Scale, the State Anxiety Scale, the Test Anxiety Scale, the Computer Anxiety Scale, and the Mathematics Class Anxiety Scale. The authors justify the naming of this first dimension by the fact that the “scales that comprise it are not related to language

behavior in a reliable manner” (p. 268). Factor 2 was named Communicative Anxiety as it obtained high loadings from French Class Anxiety, French Use Anxiety, English Class Anxiety, and the Audience Sensitivity Scale. The authors observe that “each of these measures involves, to some extent, anxiety reactions in oral communication situations” (p. 261). A further study by MacIntyre and Gardner (1991) included 19 anxiety measures, with four scales related to French L2 learning. Factor analysis provided more evidence of the differentiation between types of anxiety measures. Three factors emerged reflecting General/Social-Evaluative Anxiety, State Anxiety, and a unique Language Anxiety factor. The authors also found that the Language Anxiety factor was the only one to be related to performance on two measures of processing linguistic material in French L2.

MacIntyre and Gardner (1994) became interested in the “subtle effects” of anxiety and its sources on L1 and L2 language performance across three stages of cognitive processing: (1) language input stage, (2) processing and interpreting the language, and (3) the output stage at which knowledge of the language can be demonstrated. They developed new scales reflecting specific types of language anxiety at these three stages. The authors concluded that

[t]he potential effects of language anxiety on cognitive processing in the second language appear pervasive and may be quite subtle. Performance measures that examine only behavior at the output stage may be neglecting the influence of anxiety at earlier stages as well as ignoring the links among stages.

MacIntyre & Gardner, 1994, p. 301

The Dynamic Approach in Foreign Language (Classroom) Anxiety Research

The third phase of anxiety research, according to MacIntyre (in press), is the Dynamic Approach, which gained popularity around 2010 among SLA researchers. The aim of this approach is to situate anxiety among a range of interacting factors that affect SLA: “Anxiety is continuously interacting with a number of other learner, situational, and other factors including linguistic abilities, physiological reactions, self-related appraisals, pragmatics, interpersonal relationships, specific topics being discussed, type of setting in which people are interacting, and so on” (MacIntyre, in press, n.p.). Anxiety is seen as an emotion that is constantly fluctuating over different timescales. One study adopting this approach is Gregersen, MacIntyre, and Meza (2014), which investigated the causes of spikes in anxiety during L2 speaking. The researchers measured heart rates of six preservice teachers who were making a classroom presentation in L2 Spanish. Following the presentation, the participants met with the instructor and reviewed the videorecording of their presentation using the idiodynamic procedure (MacIntyre, 2012), which shows changes in anxiety in real time. Anxiety spikes emerged when speakers forgot words or lost the thread of their presentation. Highly anxious participants (measured with the FLCAS) were more likely to experience spikes in anxiety, possibly because they had memorized their presentations.

MacIntyre and Serroul (2015) considered the dynamic interaction of motivation and anxiety when L2 users run into lexical or grammatical difficulties. They argue that problems cascade, which they compare to four hostile horsemen. First, an inhibition system is activated by the appraisal of a clear and present threat, which shifts attention away from the language production to the interlocutor and the threat to the speaker’s positive sense of self and to the interpersonal relationship. If the difficulties persist, the speaker activates coping efforts and

starts to perceive an emerging anxiety reaction. The heightened anxiety exacerbates communication difficulties as it generates distracting, self-deprecating cognition that distracts from the communication at hand and shifts cognition toward face-saving strategies, or ways to end the communication altogether. In addition to the cognitive, emotional, and linguistic difficulties, the speaker experiences the familiar physical reactions associated with high anxiety, such as perspiration, a racing heart, shaky limbs, and butterflies in the stomach. It all leads to frustration and increased avoidance motivation, declining perceptions of competence and lower willingness to communicate (MacIntyre & Serroul, 2015). What the study shows is that “the anxiety state reflects the coalescence of a number of dynamically changing processes” (MacIntyre, in press, n.p).

Current Issues

It would be slightly depressing to state that the current issues in anxiety research in the field of the SLA are the same as before. There is some truth in this, however. It does not mean that the field has been standing still, as the previous overview clearly shows. Researchers have developed new instruments and approaches to observe the anxiety of foreign language learners and users. Demonstrating progress in science is a challenging task because it can be hard to establish clear boundaries among fields, currents, and periods, such as MacIntyre’s (in press) distinction between the Confounded, the Specialized, and the Dynamic Approaches. The complexity of anxiety research defies easy categorizations. Inevitably, approaches can overlap and coexist, and some may gain in dominance over time before losing it again. Another way of looking at the field is through a research time line such as Horwitz (2010) who identified 44 milestones “in the development of the language teaching profession’s understanding of anxiety reactions in response to L2 learning and use” (p. 154). She admits that such an exercise is inevitably subjective. The trend that she observes is quite similar to MacIntyre’s (in press) overview. Many of the early articles, Horwitz (2010) notes, “address the nature of FLA as contrasted with or related to other anxiety types [. . .] and the effects of anxiety especially on language achievement” (p. 154). Later work was more concerned “with sources of FLA and its stability or variation under different instructional or socio-cultural conditions [. . .], the relationship of FLA with other learner factors [. . .], anxieties in response to specific aspects of language learning such as listening, reading, or writing [. . .], and instructional strategies to reduce FLA” (p. 154).

Some of the old questions remain valid today, such as the negative effect of FLA/FLCA on progress in L2 development (MacIntyre, 1999; MacIntyre & Gregersen, 2012) but the reasons for asking them may have shifted over time. The questions that Elaine Horwitz, Robert Gardner, and Peter MacIntyre asked in the 1980s about the relationship between trait, state anxiety, and FLA were motivated by a desire to prove that FLA/FLCA was a unique construct. Significant relationships between other anxieties and FLA/FLCA were therefore slightly downplayed. It would not have served their call for independence of the concept by dwelling too much on its links with existing recognized forms of anxiety. They made a convincing case that FLA had both trait and state-like characteristics (MacIntyre, 2007) but that FLA was an experience that arose uniquely in foreign language classrooms or in instances of foreign language communication.

It should be noted that participants in their studies were always students who were still studying a foreign language. In other words, they were foreign language *learners* rather than experienced foreign language *users*. This distinction may seem of little importance, but I would argue that it matters. Of course, language teachers need to know about the FLA/

FLCA that their students may suffer in their classrooms, and find ways to alleviate anxiety. However, there are more foreign language users in the world than foreign language learners (cf. Cook, 2002). An exclusive focus on the emotions of children and young adults learning languages in schools and universities might create a distorted image as it ignores the majority of adult foreign language users in the world. These foreign language users are typically still developing their language skills outside school and should therefore be included within a larger ISLA context. My own research has thus generally included a wider range of ages and backgrounds of participants. As the concept of FLA/FLCA is well established in our field, we can now freely explore to what extent FLA/FLCA is linked to other personality characteristics. Finding such links poses no threat to the independence of the construct as it merely enriches our understanding of it. In fact, considerable psychological research seeks links between personality traits and various psychological dimensions.

Key Concepts

Foreign language anxiety (FLA): “The worry and negative emotional reaction aroused when learning or using a second language” (MacIntyre & Gardner, 1994, p. 27).

Foreign Language Classroom Anxiety (FLCA): “A distinct complex of self-perceptions, beliefs, feelings and behaviors related to classroom learning arising from the uniqueness of the language learning process” (Horwitz et al., 1986, p. 128).

Relationship between the anxieties of foreign language learners and users: A nested design could be imagined with Communicative Anxiety as the outer ring, with gradually smaller inner rings starting with Language Anxiety, Foreign Language Anxiety, Foreign Language Classroom Anxiety, and the anxieties linked to specific classroom activities such as speaking, listening, reading, and writing (see Figure 24.1).

Higher order personality traits: “Refer to consistent patterns in the way individuals behave, feel and think” (Pervin & Cervone, 2010, p. 228). The Big Five bipolar higher order dimensions are openness to experience, conscientiousness, extraversion versus introversion, agreeableness, and neuroticism versus emotional stability, which are situated at the summit of the hierarchy (2010, p. 228). Another higher order dimension used by some psychologists is *Psychoticism*, typified by aggressiveness and interpersonal hostility. These higher order dimensions are correlated with facets beneath them. For example, people who score high on *Openness to experience* are typically creative, original, imaginative, curious, and flexible; those at the low end of the dimension are unartistic, conservative, conventional, practical, and down to earth. People who score high on *Conscientiousness* are typically meticulous, efficient, organized, reliable, hardworking, and persevering; low scorers are typically unreliable, careless, disorganized, lazy, and negligent. *Extraverts* are typically talkative, assertive, sociable, gregarious, active, and passionate; *Introverts* tend to be shy, passive, quiet, reserved, withdrawn, and sober. People who score high on *Agreeableness* are typically friendly, good-natured, kind, trusting, cooperative, modest, and, generous; low scorers are typically cold, rude, unpleasant, critical, antagonistic, suspicious, and uncooperative. People who score high on *Neuroticism* tend to worry, to be anxious, insecure, depressed, emotional, and unstable; people at the *Emotional stability* end of the scale are typically calm, relaxed, hardy, content, even-tempered, and self-satisfied.

Distribution on personality dimensions: Scores are normally distributed, meaning that a majority of people are situated in the middle of the dimension.



Figure 24.1 Nested design of anxieties

Empirical Evidence

Higher Order Personality Traits and FLA/FLCA

Personality traits “summarize a person’s typical behavior” (Pervin & Cervone, 2010, p. 229) and psychologists agree that there are five broad, bipolar dimensions, the so-called Big Five (p. 228), which are situated at the summit of the hierarchy (for a more detailed description, see the Key Concepts box in the previous section); there are a large number of narrower facets, “lower order” personality traits, that are often correlated with Big Five traits but also explain unique variance. Trait Emotional Intelligence, for example, was shown to share more than 50% of the total variance with the Big Five personality traits (Extraversion, Neuroticism, Openness, Agreeableness, and Conscientiousness) (Petrides et al., 2010). The authors presented this overlap as a strength rather than a weakness.

The earliest empirical evidence of the link between FLCA and more general personality characteristics was already presented in the overview: Horwitz (1986) reported significant positive correlations between the FLCAS and the Fear of Negative Evaluation Scale, the Test Anxiety Scale and the Trait scale of the State-Trait Anxiety Inventory, which meant that people who are anxious in general are also typically more anxious in language learning.

MacIntyre and Charos (1996) toyed with the idea of linking language anxiety with Emotional Stability (which is the positive end of the Neuroticism dimension) in a group of Anglo-Canadian students with French L2. They noted, “individuals with lower emotional stability may be more prone to language anxiety” (p. 11). However, they decided not to investigate this possible link from emotional stability to language anxiety “because prior research has demonstrated that language anxiety is not strongly related to general trait anxiety, which would be reflected in a lack of emotional stability” (p. 11). They also found, unsurprisingly, that introverts, who are typically quieter and shy, suffered significantly more from L2 anxiety.

Dewaele (2002), in a study of 100 Belgian L1 Dutch-speaking learners of L2 French, failed to find a correlation between levels of FLA and scores on Extraversion, Neuroticism, and Psychoticism. Surprisingly, significant relationships did emerge between these three personality dimensions and the same students’ levels of FLA in L3 English: Psychoticism ($r = -.30, p < .01$), Extraversion ($r = .23, p < .05$), and Neuroticism ($r = .22, p < .05$). The effect sizes ranged from 4.8% to 9% of variance explained, which can be described as small. The hypothesis that extraverts being more talkative and optimistic would be less anxious was confirmed only for L3 English, but not for L2 French. The same puzzling finding for Psychoticism and Neuroticism defied a simple explanation. High scorers on Psychoticism were expected to be less anxious because they typically care less about being perceived positively by interlocutors, and participants scoring high on the Neuroticism scale, which reflects general trait anxiety, were expected to be more worried about their performance in *both* foreign languages, not just one. Interestingly, FLA in French turned out to be linked not to psychological variables but to social class, with students from lower social classes being significantly more anxious in French. This finding could be linked to the fact that French used to be a prestigious language in Flanders, spoken fluently by members of higher social classes. French thus used to be a social marker and this perception seemed to linger on, overriding the effects of personality traits. The finding of a relationship between personality traits and FLA for one foreign language but not for another had some unexpected implications for previous research. When Horwitz, MacIntyre, and Gardner talked about FLA and FLCA in their work, they based their findings on a *single* foreign language, and seemed to assume that relationships they uncovered would apply to all foreign languages equally. Retrospectively, it would have been interesting to investigate whether the relationships uncovered by Horwitz, MacIntyre, and Gardner over the years for the L2 also appeared in the L3 or L4 of any participants who knew more than two languages. What Dewaele (2002) showed was that interrelationships between psychological variables were more dynamic than had been assumed so far.

Dewaele (2013a) investigated the link between three global personality traits (Psychoticism, Extraversion, and Neuroticism) and levels of FLCA (Horwitz et al., 1986) in the second (L2), third (L3), and fourth (L4) languages⁴ of two groups of adult language learners and users. The first group consisted of 86 students from London, and the second group consisted of 62 students from Mallorca. All students were studying at least two foreign languages (i.e., languages learnt after the age of 3). Correlation analyses revealed a significant positive link between Neuroticism and FLCA in the L2 and L3—but not the L4—of the London

group (L2: $r = .31, p < .01$; L3: $r = .27, p < .05$; L4: $r = .30, p = .08$). Similar patterns emerged for the Mallorca group (L2: $r = .34, p < .01$; L3: $r = .50, p < .001$; L4: $r = .51, p < .01$). In other words, Neuroticism and FLCA shared between 9% and 25% of variance in most foreign languages, which can be described as small to moderate effect sizes. Psychoticism and Extraversion were unrelated to FLCA in the London group but were significantly negatively related with the L3 for the Mallorca group ($r = -.26, p < .05$ and $r = -.29, p < .05$, respectively). These are small effects sizes with 6.7% and 8.4% of variance explained. These findings further confirmed that the strength of association between personality traits and FLCA varies from language to language for the same participants, and that the effects of Extraversion and Psychoticism were inexplicably different in the two groups.

A further study involving sociobiographical variables and higher order personality traits and FLCA was that by Dewaele and Al Saraj (2015). Participants were 348 Arabic learners of English in the Arab world who filled out the Arabic Foreign Language Anxiety Questionnaire—a culturally adapted version of the FLCAS consisting of 33 items—and an Arabic version of the Multicultural Personality Questionnaire—Short Form (van der Zee, van Oudenhoven, Ponterotto, & Fietzer, 2013). Pearson correlation analyses revealed that FLCA was significantly and negatively correlated with four personality traits: Cultural Empathy ($r = -.13, p < .05$), which is strongly related to the Big Five dimension of Agreeableness; Social Initiative (strongly linked with Extraversion) ($r = -.34, p < .0001$), Openmindedness (strongly linked with Openness to experience) ($r = -.36, p < .0001$), and Emotional Stability (the positive end of the Neuroticism dimension) ($r = -.46, p < .0001$). In other words, the multicultural personality traits shared between 1.7% and 21.1% of variance with FLCA, which can be described as small to moderate effect sizes. A multiple regression analysis, including sociobiographical variables, revealed that Emotional Stability and Social Initiative together explained 18.5% of variance in FLCA, a result that is similar to the findings for Neuroticism and Extraversion in Dewaele (2013a). It thus seems that the more extravert students and the emotionally stable students—who can stay calm under “novel and stressful conditions” (van der Zee et al., 2013, p. 118)—suffered less from FLCA. The correlations between FLCA and Openmindedness and Cultural Empathy suggest that learners with an open and unprejudiced attitude toward cultural differences and an ability to empathize with the feelings, thoughts, and behaviours of culturally diverse individuals tended to suffer less from FLCA. Similar patterns emerged in Dewaele and MacIntyre (2016b). A group of 750 foreign language learners from mostly Europe and North America filled out eight items from the FLCA (Horwitz et al., 1986), the Foreign Language Enjoyment scale (Dewaele & MacIntyre, 2014) and the Multicultural Personality Questionnaire (van der Zee et al., 2013). A multiple regression analysis revealed that Emotional Stability explained 28.4% of variance in FLCA while Social Initiative explained a further 3.3% of variance. Interestingly, Cultural Empathy predicted 8% of variance of FLE.

A slightly different approach was taken by Muehlfeld, Urbig, Van Witteloostuijn, and Gargalianou (2016) who argued that gender is a crucial mediating variable between general personality traits (measured with the HEXACO Personality Inventory—Revised Version) and FLCA. The authors looked at 320 adult L1 Dutch speakers who had English as a foreign language and found that their 106 female participants experienced higher levels of FLCA (measured with a shortened version of the FLCAS), but that this association was mediated by differences in personality. The female participants scored higher on emotionality and conscientiousness—dimensions that happened to be most strongly linked with FLCA. There was a significant positive correlation between FLCA and Emotionality ($r = .34, p < .001$), which includes trait anxiety. Tests of discriminant validity did show that this trait anxiety

was psychometrically distinct from FLCA. Conscientiousness was the second personality dimension to be related to FLCA ($r = .20, p < .001$). People who score higher on this dimension tend to be well organized, dependable, and self-disciplined. The authors suggest that Conscientiousness is related to more negative and more emotional responses to speech errors. The third dimension was Extraversion ($r = -.15, p < .01$), which the authors explain by the fact that more introverted people are more likely to feel threatened by being exposed within a group. The effect sizes were thus small, explaining between 2.2% and 11.5% of shared variance.

Lower Order Personality Traits and FLA/FLCA

Research has also focused on the link between FLA/FLCA and lower order personality characteristics or constituent facets. Dewaele, Petrides, and Furnham (2008) was the first published study to link FLA with Trait Emotional Intelligence (Trait EI)—also known as emotional self-efficacy and defined as a constellation of emotional self-perceptions⁵ located at the lower (and narrower) levels of personality hierarchies. Trait EI was measured with the Trait Emotional Intelligence Questionnaire—Short Form (Petrides & Furnham, 2006). Trait EI is positively linked to Extraversion and Emotional Stability. The study considered the effects of sociobiographical variables and of Trait EI on communicative anxiety in the first language and FLA in the L2, L3, and L4 of 464 adult multilingual individuals, in five different situations (speaking with friends, colleagues, strangers, on the phone, and in public). Participants with lower levels of Trait EI suffered significantly more from FLA in almost all situations in all their languages, including their L1. Kruskal Wallis tests indicated that the effect of Trait EI was most significant in the L1 when speaking with colleagues, strangers, on the phone, and in public (all $p < .0001$). It remained significant ($p < .05$) for all situations in the L2, L3, and L4. An analysis of the χ^2 values suggest a small effect size, with Trait EI explaining between 1.7% and 4.5% of variance across languages and situations. The drop in FLA was relatively limited between the low and average Trait EI groups in the L2 and L3 but was much steeper between the average and the high Trait EI groups. One possible explanation was that the high Trait EI group had a stronger self-belief in their ability to regulate stress levels and to express themselves, and were better equipped to recognize the emotional state of their interlocutors, which led to lower levels of FLA.

These findings were confirmed in Shao, Yu, and Ji (2013) who considered the relationship between FLCA and Trait EI among 510 Chinese students in English classes. Students' scores on Trait EI and FLA ($r = .68, p < .01$) were negatively and significantly correlated with each other and explained 46% of the variance. High levels of Trait EI corresponded with low levels of FLA. Students who scored high on Trait EI and low on FLA were also found to perform better in English examinations.

Dewaele and Tsui Shan Ip (2013) looked at the effect of another psychological dimension on FLCA, a dimension that Ely (1995) had been previously linked to SLA, namely Second Language Tolerance of Ambiguity. The study was based on data from 73 secondary school students in Hong Kong, which reported on FLCA in their English classes using Horwitz et al.'s (1986) questionnaire. Results showed that students who were more tolerant of second language ambiguity were significantly less anxious in their EFL classes ($r = -.71, p < .0001$) and also felt more proficient in English. The effect size is large, as more than half of the variance is explained (50.4%). The finding was interpreted in the light of the knowledge that people feel anxious when there is ambiguity (Gudykunst, 2005), and that EFL learners in particular have to deal with ambiguity in the input, uncertainty about the exact meaning of

English words and phrases, and difficulty in recognizing unfamiliar phonemes or prosody, which raises FLCA levels. Those with lower levels of Second Language Tolerance of Ambiguity are at a particular disadvantage in that situation and will suffer more from anxiety than their peers with higher levels of Second Language Tolerance of Ambiguity.

Dewaele (in press) investigated the relationship between Foreign Language (Classroom) Anxiety and Perfectionism. Three different groups of participants provided data via online questionnaires: an international group of 58 adult multilingual English foreign language users filled out the Frost Multidimensional Perfectionism Scale (FMPS) (Frost, Marten, Lahart, & Rosenblate, 1990) and a questionnaire on Foreign Language Anxiety (Taguchi, Magid, & Papi, 2009); 69 Saudi students filled out the FMPS and the FLCAS; and 323 Japanese university students filled out the Multidimensional Self-Oriented Perfectionism Scale (Sakurai & Ohtani, 1997) and a selection of items from the FLCAS. Significant positive relationships emerged between Perfectionism and FLA/FLCA in the international group ($r = .38, p < .001$), in the Saudi group ($r = .29, p < .018$), and in the Japanese group ($r = .22, p < .0001$), suggesting that more perfectionist respondents felt more anxious when using English. The effect sizes vary from small toward moderate (ranging from 4.8% to 14.4% of variance explained). These results confirmed the findings of an earlier study by Gregersen and Horwitz (2002) who found that highly anxious participants exhibited perfectionist tendencies. Gregersen and Horwitz focused on the four most anxious and the four least anxious Chilean language students (out of a pool of 78 students who wanted to become English teachers) on the basis of the FCLAS scores. The highly anxious students were more motivated by negative than positive emotions, they delayed getting started on work that would be judged, and they perceived anything less than perfect as a failure. The authors found that the anxious learners scored significantly higher than the nonanxious learners on personal performance standards and procrastination, in other words, perfectionist tendencies.

The last two studies, by Liu and Jackson (2008) and Wang (2010) focused on Chinese learners of English. Wang (2010) looked at the effect of personality variables on FLA among 240 Chinese learners of English. The author found that learners with higher levels of English speaking anxiety scored higher on Trait anxiety ($r = .34, p < .01$) and on unwillingness to communicate with others ($r = .57, p < .01$). Higher speaking anxiety was also linked to lower rates of risk-taking in the English class ($r = -.54, p < .01$), language class sociability ($r = -.33, p < .01$), and speaking self-efficacy ($r = -.38, p < .01$). Moreover, high speaking anxiety was negatively correlated with English achievement ($r = -.36, p < .01$). The effect sizes were moderate, explaining between 10% and 32.5% of variance.

Wang's results confirmed the previous study by Liu and Jackson (2008) on 547 Chinese students of English. The authors found that FLCA (Horwitz et al., 1986) was positively correlated with unwillingness to communicate ($r = .34, p < .01$), but negatively with language class risk-taking ($r = -.46, p < .01$) and language class sociability ($r = -.35, p < .01$). The effect sizes were moderate, varying between 10% and 21% of variance explained. Further analyses showed that unwillingness to communicate and FLCA shared common predictors.

Summary and Some Epistemological and Methodological Considerations

To sum up, research has uncovered significant links between FLA/FLCA and a range of higher order personality traits (mainly Neuroticism-Emotional Stability, Introversion-Extraversion or Social Initiative, and—to a lesser extent—also Psychoticism, Conscientiousness, Openmindedness, Cultural Empathy). Similarly, relationships have been found

between FLA/FLCA and a number of lower order personality traits or psychological dimensions. These include Trait EI, Perfectionism, Trait anxiety, Unwillingness to communicate, Risk-taking in the foreign language class, foreign language class sociability and Speaking self-efficacy. The effect sizes in all studies were typically small or moderate with only a few tending toward “large” (i.e., explaining more than 36% of variance). In other words, there is no doubt that FLA/FLCA is a unique construct, but it is just one node in a large spiderweb of personality traits and states. To extend the metaphor, one could argue that the web itself is gently pushed around by the wind and by flies that may have been captured in the web. In other words, the effects of various psychological variables on levels of FLA/FLCA are not constant but dynamic and often language specific. On top of these complex interactions come other layers of sociobiographical, situational, and social variables, which could interact among themselves but also with a wide range of psychological variables. This inherent complexity has practical implications for the research designs of quantitative researchers: the number of independent variables that could have a direct or indirect effect on FLA/FLCA is so large that they cannot all be included in one massive analysis. This limitation means that quantitative researchers are forced to focus on one or two handfuls of independent variables at the most. Rather than illuminating the whole set of relationships between variables and FLA/FLCA with dazzling sunlight, they are forced to restrict themselves to particular areas with a flashlight. This narrow focus does not lessen the value of the findings but it requires intellectual honesty about their generalizability.

What this overview of research on personality and FLA/FLCA demonstrates is that we have come a long way since the early research on the good language learner. We have become aware that no single psychological characteristic can be identified as the most beneficial in SLA. We have understood that we cannot automatically generalize findings from one single context even if the statistical results allow us to reject the null-hypothesis. We have learned that individual learners cannot be isolated from their geographical, social, and historical contexts. In other words, two learners with identical psychological profiles may experience different levels of anxiety in the foreign language class and may attain very different levels of mastery in the foreign language depending on where they are in the world. The assumption that two individuals may have identical psychological profiles is problematic in itself, because their life experiences will differ: they may have fallen in love with—or started hating—different books or people from different language backgrounds; they may have spent some time abroad using the foreign language in different situations; and the period abroad may have been a happy—or a less happy—period in their life, which could have affected the perception of the language used during that time. As researchers we may search for commonality, but we need to keep in mind that unique triggers or life events may have a much bigger effect on the emotions that learners experience and on their ultimate “success” in SLA than do carefully measured dimensions (cf. Dewaele, 2013b). I realize that this situates me clearly in what MacIntyre (in press) calls the Dynamic Approach. This is fine with me, as long as it does not imply a rejection of quantification based on the argument that “SLA does not lend itself easily to quantitative investigations, because the number of confounding variables is extensive and some of them cannot be measured at the level of precision that is required” (Dörnyei, 2009, p. 242). I explained that some degree of reductionism is inevitable in quantitative research, but this does not mean that group averages “iron out idiosyncratic details that are at the heart of understanding development in dynamic systems” (Dörnyei, 2014, p. 83). Other approaches allow researchers to zoom in on idiosyncratic details. I argue that we should not discard the—by nature—incomplete

view from above for a complete view of an idiosyncratic detail. To understand the life of trees we need views from the forest as well as from individual trees. The Dynamic Approach is fine as long as it does not restrict the methods used in the exciting hunt for individual differences.

Pedagogical Implications

While foreign language learners (and users) will always have different personality profiles, and experience different levels of FLA/FLCA, teachers can do quite a lot to alleviate anxiety and boost enjoyment in their foreign language classes. Oxford (in press) explored the ideas and strategies from Positive Psychology and Abnormal Psychology to help anxious language learners change their minds. She suggested that teachers can intervene to calm learners whose language anxiety is of a social nature by allowing them to be gradually exposed to language performance situations rather than avoiding them and by using cognitive and affective techniques to face those situations. Drawing on Rational-Emotive Therapy, teachers can encourage learners to identify their negative assumptions at home, and then in a social situation forcing themselves to speak up in order to defeat the negative assumptions. Social skills training can also help learners treat their social fears. Oxford suggested that therapists or teachers can help students with high levels of generalized anxiety to identify their maladaptive assumptions and to encourage them to change their assumptions in settings that would typically trigger their anxiety. In addition to relaxation training and biofeedback teachers could help anxious learners recognize “the role of worrying and their misconceptions about worrying; having them observe their physical arousal and the triggers to their anxiety; and helping them see the world as less threatening and hence less anxiety-provoking” (Oxford, in press, n.p.). Oxford also delved in the literature on Positive Psychology and suggested that an increase in positive emotions and emotional intelligence can help learners control their language anxiety: “The learner uses ABCDE to recognise that beliefs about adversity cause consequent negative feelings (e.g., anxiety), but disputation, i.e., presenting counter-evidence, results in energisation, or a positive change of mind (Seligman, 2006)” (Oxford, in press, n.p.). Teachers can also strengthen anxious learners’ ability “to take their minds off failure or difficulties and instead visualise something interesting in the language activity or text” and help them letting go of emotional icebergs and grudges. By creating a positive classroom climate teachers can increase flow and intrinsic motivation among all learners, including the anxious ones. I joked in Dewaele (2015, p. 14) that “learners’ emotions are like wild horses (or at least, ponies). Learners can, with a little dexterity, and with a little help from teachers, harness the power of their emotions to absorb more of the FL and the culture.”

Oxford (in press) argued that anxious learners can also be encouraged to increase their agency, that is, taking responsibility for their own learning through the use of a range of cognitive, metacognitive, social, and affective strategies. Teachers can also use joking to help anxious learners overcome their negative emotions. Boosting optimism and hope among learners is also something all teachers should do. By teaching learners how to generate alternative pathways toward a particular goal and how to use positive self-talk (Oxford, 1990, 2011) teachers can help anxious students remove temporary blockages toward goals. Teachers’ adoption of an optimistic explanatory style can help learners make more positive attributions, that is, not viewing negative situations as permanent (Oxford, in press). Oxford’s conclusion is that these teacher (and therapist) interventions can help learners overcome their social or generalized anxiety.

Future Directions

There is an increasing interest in the psychology of language learning, with a first international conference on *Matters of the Mind—Psychology of Language Learning* organized by Sarah Mercer in Graz, Austria, in May 2014; a second conference *Individuals in Contexts: Psychology of Language Learning 2* organized by Paula Kalaja in Jyväskylä, Finland, in August 2016; and a third conference organized by Stephen Ryan in Tokyo in June 2018. There is room for expansion in different directions. I like the idea of looking at nonverbal language anxiety cues (cf. Gregersen, MacIntyre, & Olsen, in press), which teachers should learn to recognize.

Another avenue of investigation is the effect of type of teaching (more or less communicatively oriented) on the anxiety that learners experience. In Dewaele, Witney, Saito and Dewaele (in press), we focused on the effect of learner-internal and teacher-centred variables on self-reported levels of FLCA and Foreign Language Enjoyment (FLE) among 192 London high school students. Learner-internal variables (such as attitude toward the FL, level in the FL and gender) were found to be linked to both FLCA and FLE. Teacher-centred variables turned out to be unrelated to FLCA but strongly linked to FLE: participants reported significantly higher levels of FLE with teachers they liked, who were unpredictable, used the foreign language a lot (rather than the students' L1) and allowed sufficient time for learners to practice their oral skills.

One other way forward in research in anxiety is not to remain solely focused on this negative emotion. By bringing in positive emotions, such as FLE, into the picture, it becomes clear that mild anxiety can co-occur with enjoyment and that learners who experience more emotion overall in the foreign language classroom are more likely to progress (Dewaele & MacIntyre, 2014, 2016a; Dewaele, MacIntyre, Boudreau & Dewaele, 2016).

I strongly encourage SLA researchers to set up interdisciplinary research projects with personality, educational, cross-cultural, social, and positive psychologists. As Mercer and Ryan (2016) argue, to understand language learning psychology, we need to stretch the disciplinary boundaries. Although the present chapter was mostly focused on quantitative research, there is also a rich qualitative approach within psychology and applied linguistics that could be further explored in SLA research (see, for example, Bailey, 1983; Gkonou, in press; Tóth, 2011; Yan & Horwitz, 2008). I feel that mixed methods, combining etic and emic approaches, quantitative and qualitative methods, could contribute a lot to SLA research (Dewaele, 2013b). An exclusive focus on means, *p*-values, and variance can produce rather dry papers, yet they could be the backbone of rich and solid studies when combined with unique insights from participants, and where the voices of researchers join in duets with those of participants.

Notes

1. The authors argue: "For correlation coefficients, we suggest that *r*s close to .25 be considered small, .40 medium, and .60 large. [...] these results show very clearly that Cohen's benchmarks for small, medium, and large correlations (.1, .3, .5) underestimate and are not appropriate for interpreting those found in L2 research" (Plonsky & Oswald, 2014, p. 889). Effect sizes indicate the "magnitude of the relationship between two variables" and is calculated "by squaring a correlation estimate (*r*) with the resulting value indicating the percentage of shared variance between the two variables in question" (Loewen & Plonsky, 2016, p. 158).
2. For a more detailed analysis of the effects of FLCA, combined with foreign language enjoyment, see Dewaele et al. (2016).

3. Personal Report of Communication Apprehension (McCroskey, 1970), Fear of Negative Evaluation Scale (Watson & Friend, 1969), Test Anxiety Scale (Sarason, 1978), State-Trait Anxiety Inventory (Spielberger, 1983).
4. Defined by the chronology of acquisition.
5. Adaptability, Assertiveness, Emotion perception, Emotion expression, Emotion management (others), Emotion regulation, Impulsiveness (low), Relationships, Self-esteem, Self-motivation, Social awareness, Stress management, Trait empathy, Trait happiness, and Trait optimism.

References

- Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *Modern Language Journal*, 78, 155–168.
- Bailey, K. M. (1983). Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies. In H. W. Seliger & M. H. Long (Eds.), *Classroom-oriented research in second language acquisition* (pp. 67–103). Rowley, MA: Newbury House.
- Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning*, 25, 153–161.
- Cook, V. J. (2002). *Portraits of the L2 user*. Clevedon: Multilingual Matters.
- Dewaele, J.-M. (2002). Psychological and sociodemographic correlates of communicative anxiety in L2 and L3 production. *International Journal of Bilingualism*, 6, 23–39.
- Dewaele, J.-M. (2013a). The link between foreign language classroom anxiety and psychoticism, extraversion, and neuroticism among adult bi- and multilinguals. *Modern Language Journal*, 97, 670–684.
- Dewaele, J.-M. (2013b). Second and additional language acquisition. In Li Wei (Ed.), *Applied linguistics* (pp. 46–68). Malden, MA/Oxford: Wiley-Blackwell.
- Dewaele, J.-M. (2015). On emotions in foreign language learning and use. *Language Teacher*, 39, 13–15.
- Dewaele, J.-M. (in press). Are perfectionists more anxious foreign language learners and users? In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*. Bristol: Multilingual Matters.
- Dewaele, J.-M., & Al Saraj, T. (2015). Foreign language classroom anxiety of Arab learners of English: The effect of personality, linguistic and sociobiographical variables. *Studies in Second Language Learning and Teaching*, 5, 205–230.
- Dewaele, J.-M., & Furnham, A. (1999). Extraversion: The unloved variable in applied linguistic research. *Language Learning*, 49, 509–544.
- Dewaele, J.-M., & MacIntyre, P. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4, 237–274.
- Dewaele, J.-M., & MacIntyre, P. (2016a). Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of the language learner. In P. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215–236). Bristol: Multilingual Matters.
- Dewaele, J.-M., & MacIntyre, P. D. (2016b). *The predictive power of multicultural personality traits, learner and teacher variables on foreign language Enjoyment and Anxiety*. Paper presented at the International Conference on Language and Social Psychology, Bangkok.
- Dewaele, J.-M., MacIntyre, P. D., Boudreau, C., & Dewaele, L. (2016). Do girls have all the fun? Anxiety and enjoyment in the foreign language classroom. *Theory and Practice of Second Language Acquisition*, 2(1), 41–63.
- Dewaele, J.-M., Petrides, K. V., & Furnham, A. (2008). The effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning*, 58, 911–960.
- Dewaele, J.-M., & Tsui Shan, Ip (2013). The link between foreign language classroom anxiety, second language tolerance of ambiguity and self-rated English proficiency among Chinese learners. *Studies in Second Language Learning and Teaching*, 3, 47–66.

- Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (in press). Foreign language enjoyment and anxiety in the FL classroom: The effect of teacher and learner variables. *Language Teaching Research*.
- Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning*, 59, 230–248.
- Dörnyei, Z. (2014). Researching complex dynamic systems: “Retrodictive qualitative modelling” in the language classroom. *Language Teaching*, 47, 80–91.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the second language learner revisited*. New York: Routledge.
- Ely, C. M. (1995). Tolerance of ambiguity and the teaching of ESL. In J. M. Reid (Ed.), *Learning styles in the ESL/EFL classroom* (pp. 216–217). New York: Heinle and Heinle.
- Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research*, 14, 449–468.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.
- Gkonou, C. (in press). Towards an ecological understanding of language anxiety. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*. Bristol: Multilingual Matters.
- Gregersen, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners’ reactions to their own oral performance. *Modern Language Journal*, 86, 562–570.
- Gregersen, T., MacIntyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners foreign language anxiety. *Modern Language Journal*, 98, 574–588.
- Gregersen, T., MacIntyre, P. D., & Olsen, T. (in press). Do you see what i feel? An idiodynamic assessment of expert and peer’s reading of nonverbal language anxiety cues. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*. Bristol: Multilingual Matters.
- Gudykunst, W. B. (2005). An anxiety/Uncertainty Management (AUM) theory of strangers’ intercultural adjustment. In W. B. Gudykunst (Ed.), *Theorizing about intercultural communication* (pp. 419–457). Thousand Oaks, CA: Sage.
- Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly*, 20, 559–562.
- Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, 43, 154–167.
- Horwitz, E. K. (in press). On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*. Bristol: Multilingual Matters.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70, 125–132.
- Kleinmann, H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning*, 27, 93–107.
- Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners’ unwillingness to communicate and foreign language anxiety. *Modern Language Journal*, 92, 71–86.
- Loewen, S., & Plonsky, L. (2016). *An A–Z of applied linguistics research methods*. London/New York: Palgrave Macmillan.
- MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language teaching: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 24–45). Boston, MA: McGraw-Hill.
- MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *Modern Language Journal*, 91, 564–576.
- MacIntyre, P. D. (2012). The idiodynamic method: A closer look at the dynamics of communication traits. *Communication Research Reports*, 29, 361–367.

- MacIntyre, P. D. (in press). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*. Bristol: Multilingual Matters.
- MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology*, 15, 3–26.
- MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification. *Language Learning*, 39, 251–275.
- MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language Learning*, 41(4), 513–534.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, 44(2), 283–305.
- MacIntyre, P. D., & Gregersen, T. (2012). Affect: The role of language anxiety and other emotions in language learning. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for language learning: Insights from research, theory and practice* (pp. 103–118). Basingstoke: Palgrave Macmillan.
- MacIntyre, P. D., & Serroul, A. (2015). Motivation on a per-second timescale: Examining approach-avoidance motivation during L2 task performance. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 109–138). Bristol: Multilingual Matters.
- McCroskey, J. C. (1970). Measures of communication-bound anxiety. *Speech Monographs*, 37, 269–277.
- Mercer, S., & Ryan, S. (2016). Stretching the boundaries: Language learning psychology. *Palgrave Communications*, 2, 1–5.
- Muehlfeld, K., Urbig, D., Van Witteloostuijn, A., & Gargalianou, V. (2016). Foreign language anxiety in professional contexts: A short scale and evidence of personality and gender differences. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 17(2), 195–223. doi:10.1007/s41464-016-0007-6
- Naiman, N., Fröhlich, M., Stern, H. H., & Todesco, A. (1978). *The good language learner*. Toronto: Ontario Institute for Studies in Education.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle.
- Oxford, R. L. (2011). *Teaching and researching language learning strategies*. Harlow: Pearson Longman.
- Oxford, R. L. (in press). Anxious language learners can change their minds: Ideas and strategies from abnormal psychology and positive psychology. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications*. Bristol: Multilingual Matters.
- Pervin, L. A., & Cervone, D. (2010). *Personality: Theory and research* (11th ed.). Hoboken, NJ: John Wiley and Sons.
- Petrides, K. V., & Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology*, 36, 552–569.
- Petrides, K. V., Vernon, P. A., Schermer, J. A., Ligthart, L., Boomsma, D. I., & Veselka, L. (2010). Relationships between trait emotional intelligence and the big five in the Netherlands. *Personality and Individual Differences*, 48, 906–910.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912.
- Sakurai, S., & Ohtani, Y. (1997). Relations of “self-oriented perfectionism” to depression and hopelessness. *Shinrigaku kenkyu: Japanese Journal of Psychology*, 68, 179–186.
- Sarason, I. G. (1978). The test anxiety scale: Concept and research. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety* (Vol. 5, pp. 193–216). Washington: Hemisphere.
- Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning*, 28, 129–142.
- Seligman, M. (2006). *Learned optimism: How to change your mind and your life*. New York: Vintage.
- Shao, K., Yu, W., & Ji, Z. (2013). An exploration of Chinese EFL students’ emotional intelligence and foreign language anxiety. *Modern Language Journal*, 97, 917–929.

Jean-Marc Dewaele

- Spielberger, C. D. (1966). Theory and research on anxiety. In C. D. Spielberger (Ed.), *Anxiety and behavior* (pp. 3–20). New York: Academic Press.
- Spielberger, C. D. (1983). *Manual for the state-trait anxiety inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Swain, M., & Burnaby, B. (1976). Personality characteristics and second language learning in young children: A pilot study. *Working Papers in Bilingualism*, 11, 115–128.
- Taguchi, T., Magid, M., & Papi, M. (2009). The motivational self system among Japanese, Chinese and Iranian learners of English: A comparative study. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 66–97). Bristol: Multilingual Matters.
- Tóth, Z. (2011). Foreign language anxiety and advanced EFL learners: An interview study. *WoPaLP*, 5, 39–57.
- Tucker, R., Hamayan, E., & Genesee, F. (1976). Affective, cognitive, and social factors in second language acquisition. *Canadian Modern Language Review*, 32, 214–226.
- Van der Zee, K., van Oudenhoven, J. P., Ponterotto, J. G., & Fietzer, A. W. (2013). Multicultural personality questionnaire: Development of a short form. *Journal of Personality Assessment*, 95, 118–124.
- Wang, T. (2010). Speaking anxiety: More of a function of personality than language achievement. *Chinese Journal of Applied Linguistics*, 33, 95–109.
- Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 33(4), 448–457.
- Yan, J. X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. *Language Learning*, 58, 151–183.

L2 Instructor Individual Characteristics

Laura Gurzynski-Weiss

Background

Instructed Second Language Acquisition and Language Instructors

As the learners' expert communicatory partner, the instructor-as-interlocutor plays a critical role in directing learners' exposure and attention to language, and their access to learning opportunities in second and foreign language (L2) classrooms. Instructors design the classroom lessons and determine the nature of the opportunities learners will have to work with the target language. Likewise, instructors determine when the focus will shift to linguistic form (pre-, during, or posttask phase; at home or in-class, etc.), and if this focus will be on form in isolation (focus on forms), or within meaning-based interaction (focus on form). Instructors are the primary providers of input and feedback in L2 classrooms, and they also elicit negotiation for meaning and determine if the learner will be encouraged to incorporate that feedback immediately or later on. In Loewen's (2015) definition of instructed second language acquisition (ISLA) as a "theoretically and empirically based field of academic inquiry that aims to understand how the *systematic manipulation of the mechanisms of learning and/or the conditions under which they occur* enable or facilitate the development and acquisition of a language other than one's first" (p. 2; my emphasis), we can see the fundamental influence of the instructor in the manipulation process. Even when working within learner-centered approaches to instruction, such as task-based language teaching, L2 instructors must decide the complexity or difficulty of the tasks with which learners will engage, along with the types of pretask instruction, modeling, and pretask planning that learners will receive. Thus, even in this brief introduction it is clear that the instructor's central roles in ISLA are without question.

While decades of research have demonstrated that learners' perception and use of the aforementioned learning opportunities are influenced by their individual differences, such as age of initial exposure, motivation, working memory capacity, and anxiety, among others (e.g., Li, 2013; Mackey, Adams, Stafford, & Winke, 2010; Sheen, 2008), there has been a recent increase in studies empirically investigating the individual characteristics of

nonlearners, such as nonteaching native speakers (NSs), researchers and, most notably, the language instructor (see initial overview in Gurzynski-Weiss, 2013). As seen in the following Key Concepts box, the term *individual differences* has been largely reserved for use when discussing individual variables of learners. Following recent growth in the number of studies examining nonlearners, use of the term *individual characteristics* has been initiated (Gurzynski-Weiss, 2013) to clarify that the individual variables of a nonlearner are the focus of inquiry; in the case of this chapter, the L2 instructor.

Key Concepts

Interlocutor: An input and feedback provider for the L2 learner, often also serves as a communicative partner. Common interlocutors include instructors, nonteaching native speakers, researchers, and fellow learners. This term is often used to describe the person with more L2 experience in a communicative exchange; in this case the interlocutor's language also serves as the L2 target.

Instructor individual characteristics: Characteristics that all instructors have, and that differ according to degree or category; including but not limited to native language(s), years of teaching experience, educational background or training, engagement with research, research specialty, working memory, and sex, among others.

Learner individual differences: Characteristics that all learners have, and that differ according to degree or category; including but not limited to age, native language(s), years of study, working memory, anxiety, proficiency level, learning strategies and styles, sex, and motivation.

Nonlearner: Term used to refer to interacting individuals who are not language learners. In this chapter, the term is used in reference to language instructors, as well as researchers and expert native or near-native speakers of the target language.

This research has revealed that the provision of L2 learning opportunities varies widely between individual instructors (e.g., Gurzynski-Weiss, 2010, 2016; Lyster, Saito, & Sato, 2013), and that this instructor variance has been found to be systematic—often determined by instructors' individual characteristics. These include but are not limited to whether or not the instructors are NSs of the language they teach (e.g., Gurzynski-Weiss, 2014, 2016; Lee, Joo, Moon, & Hong, 2006; Orton, 2014) or whether they are speakers of a specific dialect (e.g., Gurzynski-Weiss, Geeslin et al., in press), their educational background or training (e.g., Gurzynski-Weiss, 2014, 2016; Mackey, Polio, & McDonough, 2004), and their years of teaching experience (e.g., Gurzynski-Weiss, 2014, 2016; Mackey et al., 2004; Wolff, van den Bogert, Jarodzka, & Boshuizen, 2014). Less investigated characteristics considered empirically include instructor engagement with research (Borg, 2010), working memory (Ziegler, in press), and research focus (A. Y. Long, in press).

In addition to individual studies, meta-analyses have also demonstrated that instructors' (and nonteaching researchers') provision of learning opportunities varies systematically and according to their individual characteristics. For example, Li's (2010) meta-analysis found interlocutor type to mediate the effectiveness of feedback (in this particular study, nonteaching NSs as compared to nonnative speaking instructors and computers). Importantly, and of particular relevance for the current chapter, these studies have appeared largely in isolation, without a unified, purposeful approach to the study of instructor individual characteristics (exceptions include Gurzynski-Weiss, 2013, 2014, 2016, in press: see also work by the AILA

ReN on Interlocutor Individual Differences in Cognition and SLA, including papers from the 2014 ReN Symposium in Brisbane and the 2015 Symposium on Interlocutor Individual Differences). The current chapter aims to cohesively present the research domain of instructor individual characteristics and will continue to argue for the necessity of approaching the study of these characteristics systematically, much like early work on learner individual differences.

Why Study Instructor Individual Characteristics?

While some may question the efforts to examine instructor characteristics, rather than focusing on language learners, as stated in the introduction of the current volume (Loewen & Sato, this volume; see also Loewen, 2015), the goal of the subfield of ISLA is to comprehensively and empirically investigate how L2 learning occurs in instructed settings. This aim necessarily includes detailed examination of all those involved—learners, of course—as well as language instructors. As stated in Gurzynski-Weiss (2013), “In order to understand the instructed L2 context thoroughly, the systematic study of instructor characteristics in relation to factors believed to mediate the success of ISLA is necessary and relevant to both linguistic theory and language teaching practice” (p. 543). After all, if L2 learners were left alone in classrooms without an instructor to select input, design and sequence tasks, provide feedback, or otherwise facilitate learning opportunities, there would presumably not be much ISLA to report. Empirical examination of the potential influence of instructor characteristics is particularly important when one considers that, while there are undoubtedly numerous language learners in any given L2 classroom with varied individual differences, the instructor’s individual characteristics have the potential to influence the learning opportunities that all learners receive.

Despite how they are often treated in ISLA research, instructed L2 learning opportunities are inherently contextualized. Figure 25.1 depicts the relationships between contextual

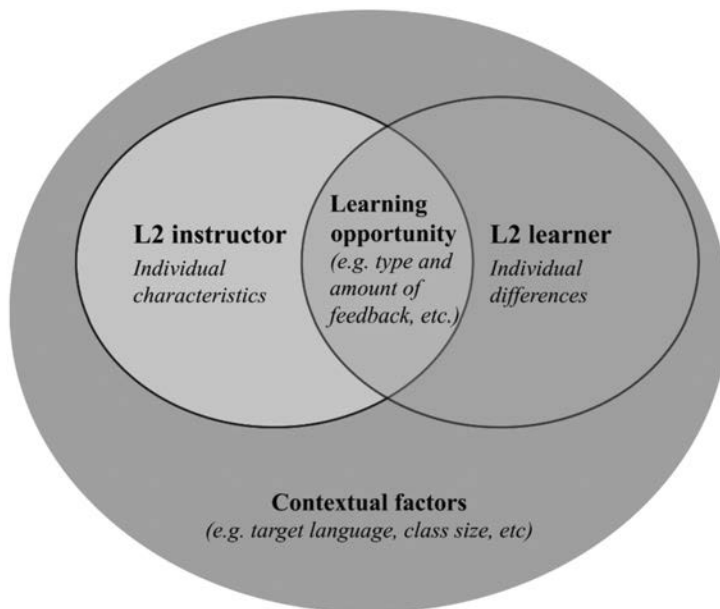


Figure 25.1 The interplay of contextual factors, learner individual differences, and instructor characteristics in ISLA

factors, learner individual differences, and instructor characteristics that are at play in L2 classrooms.

As seen within Figure 25.1, each learning opportunity (e.g., opportunities for modified output, which could also be conceptualized as dependent variables if coming from a researcher perspective) examined within instructed contexts is potentially influenced by contextual factors, such as institutional requirements, target language, logistical constraints including classroom size, time allotted for class, and the timing of a particular lesson within a given unit, among others (Gurzynski-Weiss, 2010, 2014). Learner social and cognitive individual differences, including their learning strategies (Nakatani, 2005; Zhang & Lu, 2015), motivation (Dörnyei, 2002; Dörnyei & Kormos, 2000; Hernández, 2010; Kormos & Dörnyei, 2004), and learning styles (Johnson, Prior, & Artuso, 2000; Tight, 2010), to name a few, also affect learners' attention to and use of these learning opportunities. Only recently have researchers begun to consider the long-overlooked component of instructor individual characteristics within the field of ISLA, and how they may play a critical role in determining the nature of learning opportunities provided to learners in L2 classrooms. This chapter surveys this latter area of burgeoning research, providing syntheses whenever possible, and outlining future directions as well as implications for both L2 classroom practitioners and for ISLA researchers.

Current Issues

Research examining instructor individual characteristics is expanding at a considerable pace and is the focus of active and recent discussion in the field of ISLA (see Gurzynski-Weiss, 2013, for an overview and Akbari & Dadvand, 2011; Gurzynski-Weiss, 2014, 2016; Junqueira & Kim, 2013 for empirical examples). These dialogues have focused on three principal themes: the need for theoretical grounding and expansion; the need to identify instructor individual characteristics of particular interest and relevance for ISLA theory and pedagogy; and the need to robustly operationalize each individual characteristic. Before examining what can be summarized from existing research investigating the influence of instructor individual characteristics in L2 classrooms, and to better contextualize these empirical findings, each of these current issues will be briefly addressed in turn.

Theoretical Approaches to Instructor Individual Characteristics

The majority of the research to date on instructor characteristics has been conducted within the cognitive-interactionist approach (Gass & Mackey, 2007; Hatch, 1978, 1983; Long, 1996; Schmidt, 1990, 2001; Swain, 1995, 2005) at least, for studies that state the theoretical role of the instructor (see Gurzynski-Weiss, 2014, 2016; Junqueira & Kim, 2013, etc.). The vast majority have examined instructor individual characteristics without explicitly stating the theory in which they are framed. In fact, if one examines work from the early 1980s, much of the Interaction Hypothesis and related research was inspired by discoveries that demonstrated differences in the ways learners interacted with nonlearners, such as NSs and instructors. For example, M.H. Long's earliest work (1980, 1983) examined interactional adjustments by NSs and nonnative speakers (NNSs) with language learners. Gass and Varonis (1985) examined negotiation and feedback present in NNS-NNS learner dyads and NS-NNS dyads. Perhaps the clearest example can be seen in M. H. Long's (1996)

oft-cited Interaction Hypothesis, where the role(s) of the nonlearner interlocutor are stated as theoretically central to ISLA:

Negotiation for meaning, and especially negotiation work that triggers interactional adjustments *by the NS or more competent interlocutor*, facilitates acquisition because it connects input, internal learner capacities (particularly selective attention), and output in productive ways.

pp. 451–452, my emphasis

In L2 classroom contexts, this *NS or more competent interlocutor* is inarguably the language instructor. Following the aforementioned earlier work examining different interlocutors, and the publication of the Interaction Hypothesis in 1996, attention in the field shifted to the other components within the hypothesis, most notably negotiation for meaning, learner individual differences, and output. However, as cited earlier, researchers are once again considering how the individual characteristics of nonlearners may be influencing learning opportunities, with the majority of work grounded in the cognitive-interactionist framework and focusing on L2 instructors.

Despite the trend of examining instructor individual characteristics within this interactionist approach, there is a movement for expansion into other theoretical frameworks, including investigating L2 instructor individual characteristics from a variationist perspective (Black, 2015; Geeslin, 2015; Gurzynski-Weiss, Geeslin et al., in press; Gurzynski-Weiss, Long, & Daidone, 2014; Long, Geeslin, & Gurzynski-Weiss, 2015), sociocultural perspective (Lantolf, 2015; see also Black, 2015; Shin & Choi, 2015), and through the lens of complexity theory (Larsen-Freeman, 2015; see also Mystkowska-Wiertelak & Pawlak, 2015; Serafini, 2015). Importantly, while the specific role(s) of the instructor-as-interlocutor in these frameworks differs, each of these SLA theories (among others) holds this individual as central to L2 development, and maintains that learning opportunities within L2 classrooms, particularly input and feedback, may be influenced by instructor individual characteristics. This expansion into multiple frameworks for a single topic reflects a larger cross-theoretical trend in the field, and speaks to the growing interest in instructor (and additional nonlearner, see Gurzynski-Weiss & Plonsky, in press) individual characteristic research.

Identifying Instructor Individual Characteristics of Interest

Once the theoretical role(s) of the instructor are identified within each ISLA framework, researchers can then hypothesize which instructor individual characteristics may have the most potential to differentially affect instructor provision, and learners' subsequent use, of opportunities with the L2. For instance, as seen earlier within the Interaction Approach, feedback provision and modified output opportunities are considered important learning opportunities (M. H. Long, 1996). Additionally, there may be instructor characteristics that affect learning opportunities across multiple theories. For example, one constant across SLA theories is the centrality of providing input for L2 learners. Individual characteristics that may mediate the type, amount, frequency, and contextualization of instructors' provision of input include their particular dialect, or research focus, for example, among other characteristics. In the former area, work by Gurzynski-Weiss, et al. (in press) has demonstrated that an instructor's particular Spanish dialect can influence the grammatical subject expression (whether it be explicit or null, both acceptable in Spanish) used with learners in L2 lessons. The latter variable, instructor research focus, has also been found to influence

the input instructors provide during class. For example, instructors who research phonology have been found to address pronunciation in class more frequently, while those with other research foci tend not to address pronunciation (A. Y. Long, in press). Theoretically both of these studies speak to the influence of provision of input, important across theories for ISLA. Practically, the first investigated Spanish subject expression, which, like variable structures in general, is difficult for English speakers to acquire. The second examined instructors' research background and how that influenced whether or not in-class input included explicit discussion of phonology. As language departments are heterogeneous, consisting of instructors of different research backgrounds, at least is the case for research-focused universities, having this individual characteristic potentially influencing the input (and undoubtedly additional more theory-specific concepts such as feedback and tasks) learners receive and interact with is of practical concern.

The most researched instructor characteristics to date include native language (e.g., Árvá & Medgyes, 2000; Bateman, 2008; Gurzynski-Weiss, 2010, 2016), educational background and training (e.g., Akbari & Dadvand, 2011; Gurzynski-Weiss, 2014, 2016; Junqueira & Kim, 2013; Polio, Gass, & Chapin, 2006; Tsui, 2003), and years of teaching experience (e.g., Aykel, 1997; Gatbonton, 2008; Gurzynski-Weiss, 2016; Polio et al., 2006; Tsui, 2003). These characteristics have no doubt been an initial focus because they overlap with the general education literature, are seen as both theoretically and practically relevant to the provision of learning opportunities in the L2 classroom, and are comparatively easier to operationalize than other individual characteristics such as expertise or knowledge, which may or may not correspond to years of experience or education. As Tsui (2003) has stated, 18 years of experience for one instructor may be 18 years of learning and refining skills, while for others it may be the same experience repeated 17 times (p. 13). Recently, studies have begun to focus on specific aspects within these characteristics, such as instructors' educational track (such as a master's in teaching) or research focus (culture as compared to linguistics, for example) (A. Y. Long, in press). Others have stressed the need to identify additional instructor individual characteristics of interest, such as anxiety (Tum, 2014) or working memory (Ziegler, in press). For example, Ziegler's inaugural work examining instructors' working memory in relation to their feedback provision and learner use of feedback was based on her hypothesis that greater working memory would allow instructors to provide greater amounts of feedback in the computer-mediated mode, particularly delayed feedback (after learners had completed a thought). While her empirical evidence did not corroborate this, Ziegler's findings lent support for future research to examine instructors' working memory in relationship to feedback provision in the face-to-face mode.

Operationalizing Instructor Individual Characteristics

A third trending point of discussion is the need to robustly operationalize each individual characteristic alone and in relation to other characteristics. Much like initial work on learner individual differences, research into instructor individual characteristics necessitates the examination and determination of the nature of a given individual characteristic and, in the case of certain characteristics such as teaching education and experience, among others, a determination of how the characteristic relates to or even overlaps with others. Looking to research on learner individual differences for methodological guidance, there has been considerable discussion and investigation into L2 learner anxiety, and how the types of anxiety may relate to each other—and ultimately influence SLA (e.g., Ellis, 2008; Horwitz, 2001; MacIntyre & Gardner, 1989, 1994); likewise for motivation (Dörnyei et al., 2015; Dörnyei

& Ushioda, 2009, 2013). As research on instructor characteristics continues to increase, so too must the methodological rigor.

Unfortunately, the studies that have examined instructor individual characteristics as independent variables have often failed to explicitly operationalize the instructor characteristic(s) they examine, and those that have largely rely on dichotomies to compare instructor groups by a given individual characteristic. To date, this has prohibited cross-study comparison and slowed development of individual instructor characteristics as a research domain. For example, studies that compare behaviors of “native” or “nonnative” instructors often do not operationalize what nativeness means, simply stating the L1 of the instructor, without specifying when or how the L1 was learned (e.g., Árvá & Medgyes, 2000; Cots & Díaz, 2005; Ghanem, 2015) or, alternatively, stating the country of origin of the individual instructor (e.g., Stevens, 2000; Yang, 2010). Notable exceptions include Gurzynski-Weiss (2010), who operationalized NS as having used the target language at home and/or in school more than 50% of the time in their primary (prepubescent) years, and Faez (2011), who uniquely determined NSs and NNSs by three criteria: (1) proficiency in English; (2) self-ascription as a NS or NNS; and (3) validation by others.

The use of dichotomous terminology, much like within the greater SLA field, is common throughout the domain of instructor characteristic research. For example, with the individual characteristic of teaching experience, instructors are often reduced to two categories of experienced or inexperienced (e.g., Aykel, 1997; Gatabonton, 2008; Junqueira & Kim, 2013; Mackey et al., 2004). Importantly, these two categories are operationalized very differently between studies. In Mackey et al. (2004) instructors labeled as “experienced” had been teaching 5–14 years, while in Gatabonton (2008), this same category was reserved for those with 10 or more years of experience. Polio et al. (2006) considered instructors to be experienced after 4 years, while Junqueira and Kim’s (2013) case study focused on an instructor with 20 years of experience. Additionally, as critiqued by Gurzynski-Weiss (2013, 2014), preservice or new instructors (also referred to as novice and inexperienced in the literature) are most often compared with very experienced instructors, excluding the majority of those teaching languages. The few attempts to examine characteristics in more detail have often neglected to explicitly explain the motivation behind choosing these particular categories. For example, in Shi, Wang, and Wen (2003), years of teaching experience were categorized as none (0 years), 1–4 years, or 5 or greater. Zapata and Lacorte (2007) conceptualized experience as being in one of five ranges: no experience, 1–3 years, 3–6 years, 6–10 years, or more than 10 years. Critically, no explanation was given as to how these categories were motivated, or why these specific ranges were chosen. To date there has not been a dedicated investigation or discussion as to the nature of and boundaries between individual instructor characteristics, and how they may or may not overlap. For this domain to grow and provide meaningful contribution to the field of ISLA, continued work on this specification will be a requisite next step.

An additional challenge of surveying what has been discovered thus far on instructor characteristics is the indistinguishable terminology used, which fails to distinguish research on instructor characteristics as independent variables from other studies providing this information as participant background information (Gurzynski-Weiss & Plonsky, in press; Plonsky & Gurzynski-Weiss, 2015). For example, many studies include instructors (or teachers), whether their individual characteristics are considered as independent variables or not; only recently (Gurzynski-Weiss, 2013) has the term “individual characteristics” been used to describe this research. Thus, studies examining relationships between instructor experience and their in-class error correction, for example, need to isolate experience as an independent variable, rather than simply reporting

instructor experience as background information. Clearly, there is considerable work to be done regarding the constructs of instructor individual characteristics.

Empirical Evidence

Most studies conducted to date have focused on a single instructor individual characteristic in one-shot designs, often in English-language contexts, and have focused on that characteristic in relationship to (1) instructor in-class cognitive processes and/or (2) learning opportunities provided in the L2 classroom. For example, a typical study identifies an instructor individual characteristic of interest, such as years of teaching experience, and compares a learning opportunity, like instructor elicitation of ESL student output in two-way information exchange tasks (Polio et al., 2006), in relation to the characteristic. Importantly, and perhaps appropriately at this point in this research domain, differential L2 learning outcomes have not been measured in these studies. Additionally, much like the (lack of) operationalizations of instructor characteristics, disclosure on which theoretical framework each study is grounded is unstated in the majority of the studies.

What Do We Know About Instructor Individual Characteristics in Instructed L2 Settings?

A considerable number of studies have found instructors' individual characteristics to influence their in-class cognitive processes and/or resulting behavior. Specifically, the characteristics of nativeness, teaching experience, and education/training/research background have been found to influence instructors' provision of input and feedback, two conditions for learning held as central for ISLA.

The characteristic of instructor nativeness, or whether or not the instructor is a native speaker (NS) of the language they are teaching, has been found to relate to their ability to predict vocabulary difficulties and their resulting lesson design (e.g., Reynolds-Case, 2012), the type of input they provide (e.g., Gurzynski-Weiss, et al., in press; Long, in press; Long et al., 2015; Stevens, 2000), and the amount and type of feedback they give to learners (e.g., Gurzynski-Weiss, 2010; Yang, 2010). For example, in Reynolds-Case (2012), nonnative speaker (NNS) instructors who shared their students' L1, English, were more able to predict which vocabulary words would be problematic for learners as compared to their NS counterparts, and they adjusted their lesson plans, and therefore the conditions for learning, accordingly. With respect to the type of input learners receive, instructor dialect has been found to affect whether or not a variable structure, Spanish subject expression, was provided to L2 learners (Gurzynski-Weiss, Geeslin, Long, & Daidone, in press). Stevens (2000) also examined input, specifically the /b/ sound in Spanish, and found instructor native language as well as gender and length of residence in an English-dominant society, to influence the type of input provided in L2 lessons; namely, that nonnative instructors provided /v/, which many consider to be non-target-like, significantly more than their native counterparts, as did female as compared to male instructors, and those who resided longer in the US. In terms of feedback and whether or not the instructor was a NS of the language they were teaching, Yang (2010) found NS instructors of English-as-a-foreign-language (EFL) to correct more grammatical errors, while NNSs from Taiwan focused on phonological errors; Gurzynski-Weiss (2010) also found NS instructors of Spanish FL to correct more grammatical errors compared to lexis, and that they did so more explicitly. Examining instructors' in-class feedback decisions, Gurzynski-Weiss (2016) found the characteristic of nativeness to

direct some instructors' feedback decision-making. Specifically, she found that for some NS instructors, this characteristic directed their attention to listen for errors that would impede communication with NSs who did not have experience with the students' L1 English. The attention of NNS instructors, on the other hand, was often directed to listening for errors they personally had difficulty with when they were learning L2 Spanish. Finally, one study, Lee et al. (2006), found learners' target language production to differ according to instructor NS background: NS instructors of EFL promoted learners' fluency, while NNSs (L1 Korean) promoted complexity and accuracy.

Studies examining instructors' teaching experience have found differences in the focus of their in-class cognition as well as in the amount of online reflection reported. For example, Gatbonton (2008) found instructors with 0–2 years of experience to report noticing student behavior and reactions more than instructors with more than 10 years of experience. Instructor experience has also been found to relate to how instructors describe and determine classroom management (Gurzynski-Weiss, 2010; Wolff et al., 2014), with less experienced instructors being more preoccupied with classroom management and having fewer solutions at their ready disposal. Gurzynski-Weiss (2010) also included examinations of L2 Spanish instructors' experience in relationship to in-class feedback decisions and found that less experienced instructors reflected considerably more than more experienced instructors, who tended to decide whether or not to provide feedback in a more automatized way, without reflection. In other words, the instructors reported simply responding to the error with feedback, without consciously thinking about whether or not they should correct the error, when they should correct it, or what type of feedback they should use, and so forth.

Studies have also found the characteristic of experience to influence instructor behavior in L2 classrooms. In terms of the complexity of input, Shin and Kellogg (2007) found novice instructors' input to be significantly less complex than colleagues with more than two years of experience. While Gurzynski-Weiss (2010) found relationships between Spanish FL instructors' years of experience and amount of feedback (with more experienced instructors providing more feedback), Junqueira and Kim (2013) found no feedback differences in their case study comparing an inexperienced ESL instructor with an instructor who had 25 years of experience. With respect to the type of feedback, Mackey et al. (2004) found experienced ESL instructors (4.5–15 years of experience, with a master's degree in TESOL) to provide more preemptive focus on form, recasts and explicit negative feedback than the undergraduate students who did not have teaching experience. Gurzynski-Weiss (2010) also found Spanish FL instructors with more experience to provide more explicit feedback compared to their less experienced colleagues (experience in this study was operationalized as more or less than seven years of experience). Research on instructor experience has also considered relationships between this instructor individual characteristic and learner use of learning opportunities. Specifically, Polio et al. (2006) found more experienced ESL instructors to be successful in eliciting more student output following feedback.

Instructor educational background and research focus have also been found to influence learning opportunities. With respect to the former category, Akbari and Dadvand (2011), like Gatbonton (2008) categorized instructors' reported pedagogical thoughts during lessons. Rather than teaching experience, however, a relationship was identified between education and instructors' cognitive processing. In this study, instructors with master's degrees in TESOL produced significantly more pedagogical thought units than their colleagues who had bachelor's degrees in English; there were also notable differences in instructor thought category rankings and frequencies. Gregersen (2007) also found instructor education/training to influence instructor perception of learners more than instructor experience. Examining

instructors' evaluation of students' foreign language anxiety in relationship to their teaching experience, she found brief training to play more of a role than whether or not instructors were experienced (graduate students) or inexperienced (undergraduate trainees) or were from the same country as the learners (US as compared to international). Gurzynski-Weiss (2016) also found educational background to influence instructors' in-class feedback decisions; those with SLA education (operationalized as two or more classes other than an ISLA teaching methods course) took many more factors into consideration when reflecting on learner errors and deciding to provide feedback or not. Using the same operationalization, Gurzynski-Weiss (2010) found instructors with SLA education provided different feedback—both in terms of the type (more implicit) and the amount (comparatively less)—than their colleagues whose educational focus was on literature. Additionally, instructor research background, specifically research focus on pronunciation, has been found to influence input (operationalized as instruction) that learners receive; as mentioned earlier, instructors with research expertise on pronunciation included phonology-focused instruction in their Spanish L2 classes; other instructors, despite their beliefs that this type of instruction was important, failed to provide any input (A. Y. Long, in press).

While the majority of studies thus far have focused on a single instructor individual characteristic, more recently a trend toward considering multiple characteristics within the same study can be observed. For example, Orton (2014) examined both native language (English; Chinese) and context (Australia compared to China) in relation to instructors' evaluations of L2 Chinese students' oral presentations and found an interplay between the native language and context: native Chinese-speaking Chinese instructors based in China were more likely to notice formal language features such as vowel tone in the L2 presentations, and while native English and native Chinese-speaking instructors of Chinese in Australia also noticed these formal features, they often chose to attend to the communicative side of learners' presentations. Additional studies that investigated multiple characteristics within the same study include Gurzynski-Weiss (2014), who found both research focus and teaching experience to influence graduate instructor feedback provision over consecutive semesters in Spanish L2 classrooms, and McNeill (2005), who found ESL teachers who spoke the same language as their students, as well as those with more teaching experience, to be more accurate in predicting learners' vocabulary difficulty in reading texts.

As the reader has hopefully noticed, the research trends described in the current chapter are taken, perhaps boldly, from studies where there are vast contextual differences, including target languages, immersion and traditional classroom contexts, and so forth. And while many studies were conducted in university settings within the US, others come from different countries and/or elementary levels. ISLA researchers are urged to report and consider the interplay between instructor individual characteristics, contextual factors, and learner individual differences simultaneously whenever possible. One study that has attempted this is Gurzynski-Weiss (2016), which examined instructors' native language, teaching experience, and education/research focus in relation to their in-class feedback decisions. Thirty-two L2 Spanish instructors participated in stimulated recalls, watching up to 10 feedback episodes from a 50-minute grammar lesson they taught earlier the same day. Multiple iterations of qualitative data analysis revealed that instructor characteristics filtered the instructors' attention to specific contextual factors and learner individual differences, which then led to their decision whether or not to provide feedback (along with what kind, when, and how to provide such feedback). In other words, instructor feedback decision-making was ordered, with this hierarchical nature determined by their individual characteristics (see Figure 25.2).

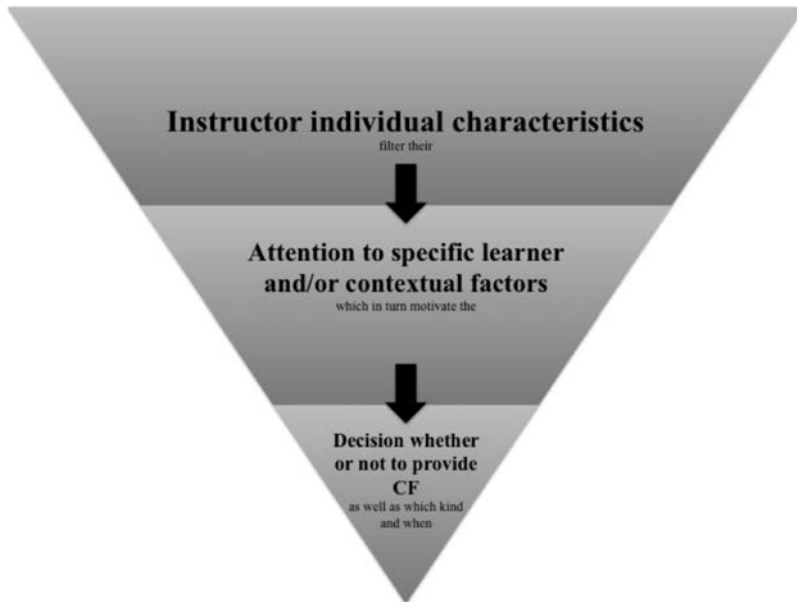


Figure 25.2 Developing a taxonomy of instructor corrective feedback decision-making

Source: Reprinted from Gurzynski-Weiss [2016] with permission from John Wiley & Sons.

Interestingly, most instructors in Gurzynski-Weiss (2016) had a dominant individual characteristic, such as educational training, while a few were consistently influenced by two or three of the individual characteristics investigated. While this one study cannot claim to be generalizable to other contexts, or even replicable within the Spanish L2 context, it does offer empirical support demonstrating that instructor characteristics, much like learner individual differences, are at play with each other. At the very least, it provides an option for moving forward in examining more than one instructor individual characteristic at a time.

As is evident from these preceding paragraphs, results thus far have been very mixed: instructor characteristics, much like learner individual differences, have been found to relate to in-class cognition and behavior. Which characteristics are most at play, if and how they systematically and reliably affect learning opportunities in L2 classrooms, and if they do so to the point of differential learning outcomes, are empirical questions still in need of more research.

Pedagogical Implications

Considering Instructor Individual Characteristics in L2 Teaching

There are several immediate pedagogical implications arising from this research. While there may not (yet) be conclusions regarding how specific individual characteristics relate to particular L2 learning opportunities, there is sufficient evidence that each individual instructor likely has characteristics that influences their in-class behavior and, ultimately, the L2 opportunities provided to learners. To incorporate this research in L2 teaching, two principal steps may be followed.

First, instructors must take inventory of their own individual characteristics and consider, and perhaps even measure, how these individual characteristics may be shaping the learning opportunities, including the input, tasks, and feedback that their students receive in the L2 classroom. Once these potential relationships between instructor individual characteristics and learning opportunities are identified, instructors may wish to ensure that their provision of learning opportunities is as balanced as possible. For example, instructors who have a research focus on morphosyntax may be predisposed to provide focused grammatical feedback considerably more than lexis, pragmatics, or pronunciation (Gurzynski-Weiss, 2014). These instructors may wish to make an effort to provide more balanced feedback to their learners by focusing on multiple areas of language. Likewise if an instructor has a particular dialect that coincides with the textbook, they could ensure their learners hear and interact with input examples of other dialects, especially those where the students may eventually study and work. This exposure is particularly important for dialects where there is variability that learners may not expect, and could even perceive as ungrammatical. Much like instructors modify lesson plans to accommodate relevant contextual factors and learners who have varying individual differences, so too could teachers balance lesson plans based on instructor individual characteristics.

Teaching Tips

- *Take inventory and identify relationships:* Instructors would do well to take inventory of their own individual characteristics and consider how these individual characteristics may be shaping the learning opportunities their students receive in the L2 classroom.
- *Make a plan to ensure balance:* Once potential relationships between instructor individual characteristics and learning opportunities are identified, instructors may wish to ensure that the types of learning opportunities they provide are as balanced as possible. Much like instructors modify lesson plans to accommodate relevant contextual factors and learners who have varying individual differences, so too may we balance lesson plans based on instructor individual characteristics.

It is important to mention that evaluation based on instructor characteristics is not part of this research. In other words, there is no goal, explicit or otherwise, to examine which instructor characteristics are “better” than others. Given that students have multiple and diverse instructors over the course of their L2 studies, there exists the possibility that instructor individual differences may not have a lasting impact on learners’ ISLA, even if there are measureable differences that influence learning opportunities within a given semester. However, this is an empirical question that must be answered via research once there is a more robust understanding of the nature of each instructor individual characteristic, as discussed in the following section.

Future Directions

In addition to the ongoing work on identifying, operationalizing, and robustly measuring theoretically and practically motivated instructor individual characteristics as outlined earlier, there are several ways future studies on instructor individual characteristics can learn from the existing research and contribute most meaningfully to the larger ISLA field. First,

we must come to an empirically grounded consensus on the most appropriate operationalizations for each instructor individual characteristic. Much like work completed on learner individual differences, such research necessitates examination and determination of the nature of a given individual characteristic and, in the case of certain characteristics, theoretical and empirical investigation of how the characteristics relate to others or even overlap. Likewise, we must conceptualize how instructor individual characteristics may go beyond simple dichotomies whenever possible. A NS instructor, for example, is not the polar opposite of a NNS instructor, nor does the characteristic assignation identically signify across individuals. This need to operationalize works in tandem with the need to use common terminology across studies. For instance, referring to research examining instructor background variables as instructor individual characteristic research would be impactful and greatly facilitate larger discussion across studies; at the very least it would permit electronic searches for synthesis. Alongside the theoretical considerations of where one individual characteristic ends and another begins, we must determine how to best measure each instructor individual characteristic, and validate these measurements, in order to be able to determine which characteristics are stable (e.g., sex) and which change over time (e.g., teaching experience or educational background for graduate student instructors), as well as why, how, and what this means for the ISLA context.

Future studies will need to measure relationships between instructor individual characteristics, learner individual differences, and learning opportunities and of course, in time, if there are links between instructor individual characteristics and differential learning outcomes, if this occurs across proficiency levels, or if there is a decrease in influence of individual characteristics once learners reach higher proficiency levels, as has been found to occur with learner individual differences (Geeslin, Linford, Fafulas, Long, & Diaz-Campos, 2013). Additionally, much of the existing research has investigated a single instructor individual characteristic, the vast majority in English language contexts, and often in one-shot designs. Future research must conduct in-depth case studies, particularly descriptive research in non-English contexts and with additional L2s, as well as conduct larger studies to see if results corroborate across contexts, and to provide a more complete picture of the many factors involved in ISLA.

Conclusions

The current chapter presented *instructor individual characteristics*, an important component within ISLA research and pedagogy. Highlighting the research that has been undertaken thus far, and recent calls for ISLA to thoroughly examine all aspects of instructed L2 contexts, the chapter argued for the need to consider instructor individual characteristics as part of the multifaceted L2 classroom, alongside contextual factors and learner individual differences. L2 instructors were urged to take inventory of their own individual characteristics, examine how these characteristics may influence the learning opportunities present in their own classrooms, and work to ensure balance, just as instructors are encouraged to do for contextual factors and learner individual differences. ISLA researchers were challenged to examine the role of the instructor in their own work, and how instructor individual characteristics may work as mediating or moderating factors within their own datasets, whether this examination occurs in ongoing projects or potential reanalysis of published work. As argued throughout the chapter, in order to thoroughly understand the ISLA environment, we must consider all interlocutors present in instructed settings, and that each of these interlocutors are in fact individuals, with their own unique set of characteristics. It is the author's hope that this chapter

Laura Gurzynski-Weiss

inspires other ISLA researchers to examine this multifaceted nature of instructor individual differences within instructed contexts. Only then will we be able to comprehensively and confidently understand how SLA occurs in instructed settings.

References

- AILA ReN on Interlocutor & Instructor Individual Differences in Cognition and Second Language Acquisition [website]. Retrieved from <http://www.individualdifferencesinsla.com/>
- Akbari, R., & Dadvand, B. (2011). Does formal teacher education make a difference? A comparison of pedagogical thought units of B.A. versus M.A. teachers. *Modern Language Journal*, 95, 44–60.
- Árva, V., & Medgyes, P. (2000). Native and non-native teachers in the classroom. *System*, 28, 355–372.
- Aykel, A. (1997). Experienced and student EFL teachers' instructional thoughts and actions. *Canadian Modern Language Review*, 53(4), 677–704.
- Bateman, B. (2008). Student teachers' attitudes and beliefs about using the target language in the classroom. *Foreign Language Annals*, 41(1), 11–28.
- Black, M. (2015, October). Interlocutor L1 effects on L2 sociolinguistic competence: A variationist analysis of L2 French *ne*-deletion. *Variationist approach*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Borg, S. (2010). Language teacher researcher engagement. *Language Teaching*, 43, 391–429.
- Cots, J., & Díaz, J. (2005). Constructing social relationships and linguistic knowledge through non-native-speaking teacher talk. In E. Llurda (Ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession* (pp. 85–105). New York: Springer.
- Dörnyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 137–158). Amsterdam/Philadelphia: John Benjamins.
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4(3), 275–300.
- Dörnyei, Z., MacIntyre, P., & Henry, A. (Eds.). (2015). *Motivational dynamics in language learning*. Bristol: Multilingual Matters.
- Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*. Bristol: Multilingual Matters.
- Dörnyei, Z., & Ushioda, E. (2013). *Teaching and researching motivation* (2nd ed.). New York: Routledge.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Faez, F. (2011). Reconceptualizing the native/nonnative speaker dichotomy. *Journal of Language, Identity and Education*, 10(4), 231–249.
- Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175–199). New York: Routledge.
- Gass, S. M., & Varonis, E. M. (1985). Variation in native speaker speech modification to nonnative speakers. *Studies in Second Language Acquisition*, 7, 283–302.
- Gatbonton, E. (2008). Looking beyond teachers' classroom behavior: Novice and experienced ESL teachers' pedagogical knowledge. *Language Teaching Research*, 12(2), 161–182.
- Geslin, K. L. (2015, October). *Variationist approach*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Geslin, K. L., Linford, B., Fafulas, S., Long, A. Y., & Díaz-Campos, M. (2013). The L2 development of subject form variation in Spanish: The individual vs. the group. In J. Cabrelli Amaro, G. Lord, A. de Prada Pérez, & J. E. Aaron (Eds.), *Selected proceedings of the 16th Hispanic Linguistics Symposium* (pp. 156–174). Somerville, MA: Cascadia Proceedings Project.
- Ghanem, C. (2015). Teaching in the foreign language classroom: How being a native or non-native speaker of German influences culture teaching. *Language Teaching Research*, 19(2), 169–186.

- Gregersen, T. (2007). Breaking the code of silence: A study of teachers' nonverbal decoding accuracy of foreign language anxiety. *Language Teaching Research*, 11(2), 209–221.
- Gurzynski-Weiss, L. (2010). *Factors influencing oral corrective feedback provision in the Spanish foreign language classroom: Investigating instructor native/nonnative speaker status, SLA education, and teaching experience* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.
- Gurzynski-Weiss, L. (2013). Instructor characteristics and classroom-based SLA of Spanish. In K. L. Geeslin (Ed.), *The handbook of Spanish second language acquisition* (pp. 530–546). Malden, MA: Wiley-Blackwell.
- Gurzynski-Weiss, L. (2014). Graduate instructor in-class cognition and feedback provision over time. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. M., Miguel, A. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected proceedings of the 2012 Second Language Research Forum*, 227–239.
- Gurzynski-Weiss, L. (2016). Factors influencing Spanish instructors' in-class feedback decisions. *Modern Language Journal*, 100(1), 255–275.
- Gurzynski-Weiss, L. (Ed.). (in press). *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors*. Amsterdam/Philadelphia: John Benjamins.
- Gurzynski-Weiss, L., Geeslin, K. L., Daidone, D., Linford, B., Long, A. Y., Michalski, I., & Solon, M. (in press). L2 classrooms as multifaceted sources of input: The synergy of variationist and usage-based approaches. In A. Tyler, L. Ortega, M. Uno, & H. I. Park, (Eds.) (in press). *Usage-inspired L2 instruction: Researched pedagogy*. Amsterdam: John Benjamins.
- Gurzynski-Weiss, L., Long, A. Y., & Daidone, D. (2014, March). Instructor oral input in Spanish L2 classrooms: Examining patterns of subject expression use. In L. Gurzynski-Weiss & K. L. Geeslin (Co-organizers), *Widening the lens on variation in and across multi-faceted sources of input*. Colloquium conducted at Georgetown University Roundtable on Linguistics, Washington, DC.
- Gurzynski-Weiss, L., Geeslin, K. L., Long, A. Y., & Daidone, D. (in press). Instructor oral input in Spanish L2 classrooms: Examining patterns of subject expression use. To appear in L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors*. Amsterdam/Philadelphia: John Benjamins.
- Gurzynski-Weiss, L., & Plonsky, L. (in press). Look who's interacting: A scoping review of research involving non-teacher/non-peer interlocutors. In L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors*. Amsterdam/Philadelphia: John Benjamins.
- Hatch, E. (1978). Discourse analysis and second language acquisition. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 401–435). Rowley, MA: Newbury House.
- Hatch, E. (1983). *Psycholinguistics: A second language perspective*. Rowley, MA: Newbury House.
- Hernández, T.A. (2010). The relationship among motivation, interaction, and the development of second language oral proficiency in a study-abroad context. *Modern Language Journal*, 94(4), 600–617.
- Horwitz, E. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, 21, 112–126.
- Johnson, J., Prior, S., & Artuso, M. (2000). Field dependence as a factor in second language communicative production. *Language Learning*, 50, 529–567.
- Junqueira, L., & Kim, Y. (2013). Exploring the relationship between training, beliefs, and teachers' corrective feedback practices: A case study of a novice and an experienced ESL teacher. *Canadian Modern Language Review*, 69(2), 181–206.
- Kormos, J., & Dörnyei, Z. (2004). The interaction of linguistic and motivational variables in second language task performance. *Zeitschrift für interkulturellen Fremdsprachenunterricht*, 9(2), 1–17.
- Lantolf, J. (2015, October). *Sociocultural theory*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Larsen-Freeman, D. (2015, October). *Complexity theory*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.

- Lee, H., Joo, K., Moon, J., & Hong, Y. (2006). The effects of interlocutor type on learners' output production: NS teacher, NNS teacher and NNS peers. *Korean Journal of Applied Linguistics*, 22(2), 201–220.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60(2), 309–365.
- Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal*, 97(3), 634–654.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Long, A. Y. (in press). Investigating the relationship between instructor research background and pronunciation-related instruction and oral corrective feedback. To appear in L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors*. Amsterdam/Philadelphia: John Benjamins.
- Long, A. Y., Geeslin, K. L., & Gurzynski-Weiss, L. (2015, October). Examining the role of instructor first language in classroom-based oral input. *Variationist approach*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Long, M. H. (1980). *Input, interaction and second language acquisition* (Unpublished doctoral dissertation). University of California, Los Angeles.
- Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, 5, 177–193.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40.
- MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification. *Language Learning*, 39(2), 251–275.
- MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of induced anxiety on cognitive Processing in the second language. *Language Learning*, 44, 283–305.
- Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60(3), 501–533.
- Mackey, A., Polio, C., & McDonough, K. (2004). The relationship between experience, education, and teachers' use of incidental focus-on-form techniques. *Language Teaching Research*, 8, 301–327.
- McNeill, A. (2005). Non-native speaker teachers and awareness of lexical difficulty in pedagogical texts. In E. Llurda (Ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession* (pp. 107–128). New York: Springer.
- Mystkowska-Wiertelak, A., & Pawlak, M. (2015, October). The dynamic nature of willingness to communicate: Exploring the impact of interlocutor variables. *Complexity theory*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *Modern Language Journal*, 89, 76–91.
- Orton, J. (2014). Comparing teachers' judgments of learners' speech in Chinese as a foreign language. *Foreign Language Annals*, 47, 507–526.
- Plonsky, L., & Gurzynski-Weiss, L. (2015, October). *Interlocutor individual differences: A methodological perspective and scoping review*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Polio, C., Gass, S. M., & Chapin, L. (2006). Using stimulated recall to investigate native speaker perceptions in native-nonnative speaker interaction. *Studies in Second Language Acquisition*, 28, 237–267.
- Reynolds-Case, A. (2012). Exploring how non-native teachers can use commonalities with students to teach the target language. *Hispania*, 95(3), 523–537.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

- Serafini, E. J. (2015, October). Applying a dynamic systems lens to interpreting the interaction between cognitive and psychosocial resources in adult second language learners. *Complexity theory*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, 58(4), 835–874.
- Shi, L., Wang, W., & Wen, Q. (2003). Teaching experience and evaluation of second language students' writing. *Canadian Journal of Applied Linguistics*, 6, 219–236.
- Shin, J. E., & Kellogg, D. (2007). The novice, the native, and the nature of language teacher expertise. *International Journal of Applied Linguistics*, 17(2), 159–177.
- Shin, S.-Y., & Choi, J. (2015, October). Examining sociocultural determinants of second language proficiency using structural equation modeling. *Sociocultural theory*. Symposium on Interlocutor Individual Differences, Indiana University, Bloomington, IN.
- Stevens, J. J. (2000). On the labiodental pronunciation of Spanish /b/ among teachers of Spanish as a second language. *Hispania*, 83(1), 139–149.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 471–484). Mahwah, NJ: Lawrence Erlbaum.
- Tight, D. G. (2010). Perceptual learning style matching and L2 vocabulary acquisition. *Language Learning*, 60, 792–833.
- Tsui, A. (2003). *Understanding expertise in teaching: Case studies of ESL teachers*. Cambridge: Cambridge University Press.
- Tum, D. O. (2014). Foreign language anxiety's forgotten study: The case of the anxious preservice teacher. *TESOL Quarterly*, 49(4), 627–658.
- Wolff, C. E., van den Bogert, N., Jarodzka, H., & Boshuizen, H. P. (2014). Keeping an eye on learning differences between expert and novice teachers' representations of classroom management events. *Journal of Teacher Education*, 66(1), 68–85.
- Yang, Y. C. C. (2010). *Recasts in the EFL classroom: A comparison of native and nonnative teachers* (Unpublished doctoral dissertation). University of Texas, Austin.
- Zapata, G. C., & Lacorte, M. (2007). Preservice and inservice instructors' metaphorical constructions of second language teachers. *Foreign Language Annals*, 40(3), 521–534.
- Zhang, X., & Lu, X. (2015). The relationship between vocabulary learning strategies and breadth and depth of vocabulary knowledge. *Modern Language Journal*, 99(4), 740–753.
- Ziegler, N. (in press). Instructor individual differences: A first look at working memory, feedback, and modified output opportunities. To appear in L. Gurzynski-Weiss (Ed.), *Expanding individual difference research in the interaction approach: Investigating learners, instructors, and other interlocutors*. Amsterdam/Philadelphia: John Benjamins.

Child ISLA

*Rhonda Oliver, Bich Nguyen,
and Masatoshi Sato*

Background

While Second Language Acquisition (SLA) emerged as a discipline in the 1960s (Montrul, 2006; Ortega, 2001), Child SLA inquiry did not begin in earnest until the following decade (Dixon et al., 2012) with studies such as those by Dulay and Burt (1974) (discussed in detail later in this chapter), and then by Huang and Hatch (1978). Child instructed SLA (ISLA), defined as the evolving, gradual, and dynamic growth of children's second language (L2) occurring primarily in the classroom and facilitated by the support of teachers (Nicholas & Lightbown, 2008; Spada & Lightbown, 2008), continues to be far less researched than other areas of ISLA (Foster-Cohen, 2010; Montrul, 2004; Pica, 2005; Simon, 2010; Spada, 2015). We begin this chapter by addressing the theoretical and methodological reasons for the comparatively slower development of this research area, namely that: (1) Child ISLA has been overshadowed by vibrant first language (L1) acquisition, adolescent SLA, and adult SLA research; (2) Child ISLA is a particularly challenging area in that the L2 child's language and sociocognitive behaviour are not as entrenched as that of an L2 adult, resulting in considerable individual linguistic variability; (3) data from Child ISLA have been used as external evidence to consolidate existing linguistic theories, but have rarely been used to develop new linguistic theories; (4) ethical issues present particular difficulties for those working with children; and (5) undertaking research with children potentially can be more time-consuming in nature than working with adults (e.g., the need to develop rapport to ensure accurate responses and children's difficulty with engaging for long periods of time mean frequent and repeated data collections, their level of "distractability" requires careful and considered materials development).

Researchers investigating Child SLA have revealed important differences between SLA and L1 acquisition, and between child and adult SLA processes and products, particularly in relation to the age of onset (or age of acquisition: AOA), the amount of exposure, accuracy orders, cross-linguistic influence, developmental sequences, and so on. In the following section, we provide empirical evidence of Child ISLA by reviewing these areas, including the role of interaction and its mediating variables, and, of most relevance

to ISLA, the way in which attention leads to L2 acquisition, particularly in classrooms where the effectiveness of form-focused instruction and corrective feedback have been examined. We conclude this chapter by providing pedagogical implications based on the updated research findings.

Current Issues

Definition of Child L2 Learners

While research indicates that Child SLA is different from adult SLA primarily due to the age at which acquisition begins, operationalization of *child* has differed depending on researchers. From a generative perspective, Haznedar and Gavruseva (2008) define child L2 learners as those who have acquired the fundamentals of the L1 and have onset exposure to the L2 between the ages of 4 and 8. In this way they also draw a distinction between Child SLA and simultaneous bilingualism, where acquisition of two languages occurs simultaneously since birth. Ionin (2008) supports this categorisation, arguing that acquisition that occurs after this period intensifies the possibility of L1-transfer. In her study, Ionin compared the acquisition of aspectual morphology between older children aged 8–9 with younger ones aged 6–7 and revealed that the former group exhibited more L1-related semantic errors. In a similar way Nicholas and Lightbown (2008) observe child learners can be distinguished as being either younger learners (aged 2–7) or older learners (8–13) with the distinction based on the emergence of literacy at around the age of 7 and on differences in ultimate attainment. Note, for these authors acquisition of two languages before the age of two is considered to be simultaneous bilingualism.

The decline of ultimate attainment in SLA depends on the age acquisition begins (i.e., age of onset); however, this decline is gradual, not occurring at a certain age, and perhaps for this reason, Child SLA researchers have not specified a definite year level at which to include or exclude child participants. In addition, and somewhat surprisingly given its importance, most Child ISLA studies do not include information about the age of onset. Hence, in this chapter, we will be inclusive and report on those age groups between preschool years to around the beginning of secondary schooling (i.e., from 2 to 14 years old). We do acknowledge, however, that the process of acquisition and the effectiveness of instruction appear to be mediated by the age of the child learners. In the following sections we discuss findings relevant to these age effects and how these have influenced theory development and methodology of Child SLA research.

Plasticity Versus Entrenchment

An L2 child's language, emotional development, and sociocognitive behaviour are not as entrenched as that of an L2 adult (Simon, 2010). For example, Park (2014) reports a mixed result with respect to L2 Korean children's adherence to the English principle of Given-before-New (i.e., a known discourse entity, or the given, always precedes a new discourse entity, or the unknown), whereas the L2 adults in the study appeared to align more consistently with their default L1 preference of New-before-Given, showing a distinct direction of transferability and together highlighting clear differences between children and adults. In brain research, imaging data during brain activation episodes indicate that the sensory cortex is more plastic in early stages of life and becomes less so as a person reaches adulthood, which has a significant impact on their perception and language learning (Shibata,

Rhonda Oliver et al.

Watanabe, Sasaki, & Kawato, 2011). These instabilities add particular difficulties to interpretations and generalizations of Child SLA studies.

Development of New Linguistic Theories

Due to the developmental stages of Child SLA described earlier, data from Child ISLA have been used as external evidence to either consolidate or validate existing linguistic theories, but have rarely been used to develop new linguistic theories (Simon, 2010). According to Cook (2010), Child ISLA research often involves the analysis of the subjects' writing or speech in "a bottom-up data-led process rather than a top-down theory-based one" (p. 137). This observation appears to hold true for most studies reviewed in this chapter (e.g., Geva & Yaghoub Zadeh, 2006; Rocca, 2007; Unsworth, 2007). For example, Lakshmanan (1994) explores how four young children (two L1 Spanish, one L1 Japanese, and one L1 French) acquire English null subjects (subject omissions) and morphological uniformity (e.g., using a copula and auxiliaries *be*, *have*, and *do* without inflectional changes) to validate Chomsky's Universal Grammar, a theory which claims that all children learn human language the same way regardless of their linguistic, cultural, or educational background.

Methodological Challenges

There are considerable methodological challenges surrounding Child ISLA research. Ethical issues present particular difficulties for those working with younger learners. For instance, obtaining permission to undertake research with minors can be complex and fraught. Geva and Zadeh (2006) documented the complexity of obtaining consent from their L2 child participants' parents or guardians. First, the consent form had to be written in two languages, English and the child's L1. Given the diversity of the children's L1s (Cantonese, Punjabi, Tamil, and Portuguese), this was deemed a remarkably daunting task, but if written agreement was not provided, data could not have been collected, hence reducing the sample size. In another study, conducted within the framework of participatory research, Pinter and Zandian (2015) reported that even after informed consent was obtained from their parents, and assent given by 10- and 11-year-old participants, and despite the explanations given prior to the commencement of the study, it was clear that the children did not understand all the information provided, as a number expressed surprise about the study at the poststudy interview.

In addition, the "time-consuming nature of research with children" (Pinter, Kuchah, & Smith, 2013, p. 486) presents a considerable obstacle for Child ISLA studies. In fact, most Child ISLA experimental studies involve the quantification of linguistic growth and change, which can only be done over an extended period of time. For example, in Chilla, Haberzettl, and Wulff's (2013) study, the children were videotaped once a month over 4 years. In the Spanish school context, Muñoz (2006) investigated L2 children's literacy development after 200 hours, 416 hours, and 726 hours of instruction, while Sollars and Pumfrey's (1999) study, also conducted in Spain, followed their child participants from the time they were in year 1 until they commenced year 3. However, it is not just the ethnographic and longitudinal nature of Child ISLA that makes such research a "time-consuming enterprise" (Spyrou, 2011, p. 18), it is also because there is an important need to meet the child participants a number of times, many more than for adult participants, to build trust and rapport before children open up to interviews and act normally in classroom observations (e.g., Pinter & Zandian, 2015).

Empirical Evidence

Child ISLA researchers have investigated a range of issues, including the similarities and differences between L1 and L2 child acquisition, age of onset or AOA, the role of interaction and its mediating variables, and the role of attention, particularly in the context of Child ISLA. In this section, we will overview those issues by drawing on age differences in the route of acquisition according to exposure, accuracy orders, cross-linguistic influence, developmental sequences, and so on.

Key Concepts

Child ISLA: Child instructed second language acquisition occurs primarily in the classroom with the support of teachers. It is an evolving and gradual process, reflecting dynamic growth of children's second language.

Age of acquisition (AOA): The age at which the L2 learning begins. It is alternatively described as the age of onset.

Maturation constraints: Physiological and cognitive factors, increasing with age, that appear to impact on language acquisition.

Ultimate attainment: The eventual level of language proficiency attained by an individual language learner.

Child SLA Versus First Language Acquisition

The small but significant body of research that involves the comparison between L1 and L2 child acquisition includes work to uncover the similarities and differences in their rate and ultimate attainment, as well as work to unveil the cognitive mechanisms underpinning acquisition. For example, based on the seminal work of Dulay and Burt (1974), who examined the natural sequences of language acquisition of L2 Spanish and Chinese children compared with L1 English children, Rocca (2007) examined L1 English children and L1 Italian children's acquisition of L2 tense-aspect (i.e., Italian and English respectively). The six child participants aged 7–8 attended language schools in either Italy or England. The bidirectional research found that L2 child acquisition is distinct from L1 child acquisition because: (1) the L2 English children appeared to systematically overextend the progressive to stative verbs whereas the L1 English children primarily used the progressive morpheme for activities and only occasionally for states; (2) the L2 Italian children appeared to overproduce the progressive aspect, underproduce the perfect aspect, and overgeneralize the perfective auxiliary—patterns that have been rarely observed in a range of L1 acquisition studies.

Other studies have, however, reported that child L1 acquisition and ISLA share more commonalities than differences. For example, Unsworth (2007) compared L1 and L2 Dutch child acquisition. The two groups of participants were of a comparable age at the time of research (i.e., 7–13 years). The L2 children were English speakers who learned Dutch as a second language at an international school. Each group was given two tasks: (1) a production task that required them to produce target forms of nonscrambled and scrambled direct objects (i.e., movement of the object to an adjoined position), and (2) an interpretation task concerning the forms used to describe the objects. The results of each group's developmental progression were then calculated separately. Generally the study showed that, contrary to expectations, both L1 and L2 child participants demonstrated a significant discrepancy

Rhonda Oliver et al.

between their production and comprehension of the scrambled and nonscrambled objects with both groups generally exhibiting more advanced production than comprehension of these features. The finding suggests that children (be they L1 or L2) find it more difficult to understand other people's meaning than to convey their own meaning.

Similarities have also been found in other developmental aspects. For example, Geva and Yaghouz Zadeh (2006) examined the reading efficiency of 181 ESL children (mean age = 7.3) and L1 children (mean age = 7) attending English speaking schools in Canada. A series of cognitive, linguistic and reading measures were administered to the participants. The cognitive and linguistic tasks comprised: (1) a nonverbal intelligence test in which the participants were asked to complete four subtests: pattern completion, reasoning by analogy, serial reasoning, and spatial visualization; (2) a rapid automatized naming (RAN) task; (3) a phonological awareness exercise that required the isolation and deletion of phonemes; (4) a picture vocabulary test that asked the children to provide one-word labels to pictures; and (5) an aural grammatical judgment test. In addition to these tasks, the participants also engaged in the following reading exercises: (1) a word attack test that required the decoding of pseudo-words; (2) a word recognition task where the children had to read 42 unrelated words; (3) Reading Efficiency Measures that measured ability to read letters and words quickly, and to use contextual clues for word identification; and (4) a word efficiency task that required the children to read two narrative texts. Quantitative analysis of the measures indicated that both groups exhibited similar results on cognitive tasks and reading exercises.

In contrast to the findings of Chilla et al. (2013), Rocca (2007) provided further evidence for similarities between L1 and L2 acquisition. Specifically, she found no major differences between L1 and young L2 German children's use of auxiliaries based on analysis of longitudinal data obtained from the CHILDES corpus. Three L1 German children and seven L2 German children, with an AOA of between 3 and 7, had their speech audiotaped once a month, over a period of 4 years, starting from when they entered school, where they received regular L2 input and were obliged to produce the L2 frequently. By examining sentences with overt subjects, thematic verbs and auxiliaries, the study reported that both groups employed similar "placeholder strategies" (e.g., use of a dummy verb that is semantically empty such as *doen* [do]). Also drawing on corpus data, a similar conclusion was reported by Cornips (2013) with regard to L1 and L2 children's use of Dutch auxiliaries *gaan* (go) and *doen* (do). The findings of these classroom based studies align with those reported in naturalistic settings (e.g., at home or in the playground); furthermore, the similarities between L1 and L2 acquisition are more apparent where linguistic errors are concerned: L2 children's errors are similar to those made by their L1 peers (see Gass & Selinker, 2008; Spada & Lightbown, 2010; Lightbown & Spada, 2013; Spada, 2015 for a detailed review of relevant naturalistic studies). L2 children also exhibit similar sequences of morpho-syntactic acquisition as their L1 peers, namely morpheme and phrasal acquisition and their developmental sequences (see Gass & Selinker, 2008; Iwasaki, 2008; Larsen-Freeman & Long, 1991; Lightbown & Spada, 2013; Spada & Lightbown, 2010).

Age of Acquisition

A second area of Child ISLA relates to the timing of when a learner begins learning the L2. Motivated in part by the critical period hypothesis (more recently referred to as the sensitive period), a theory that concerns an optimal age range for L2 acquisition (Lenneberg, 1967; Penfield & Roberts, 1959), and by maturational constraints, such as physiological

and cognitive factors that allegedly impinge on language growth (Long, 1990), a number of researchers have compared the way acquisition occurs, and also the rate and ultimate attainment achieved by younger and older L2 child learners and in comparison with L2 adult learners.

Many of the findings show that despite common belief, older learners acquire their L2 more effectively than their younger counterparts, at least in the short to medium term. For instance, Muñoz (2006) reported that older EFL children (aged 11) acquired morphology, syntax, and literacy-based skills faster than younger EFL children (aged 8) in Spanish school settings due to advantages in their cognitive development (i.e., mature growth of their brain structure and organisation). The two groups, drawn from 30 state schools, were exposed to the same number of instructional hours in English. Data were collected three times: after 200 hours, 416 hours, and 726 hours of instruction. At each interval, the participants were given an extensive test battery that included dictation, a cloze test, a listening comprehension test, grammar exercises, a composition, an oral narrative, an oral interview, phonetic imitation, phonetic discrimination, and role play to assess their four macro-skills (speaking, listening, reading, and writing). The comparative analyses over time indicated that the scores were significantly higher in the 11-year-old group, especially in literacy-oriented tests such as grammar and writing, suggesting “older L2 learners have a maturational advantage over younger L2 learners in academic tasks, in accordance with their superior cognitive development” (Muñoz, 2006, p. 4).

Similarly, Sollars and Pumfrey (1999), who studied 156 primary EFL children in Malta (mean age = 5.06 years, with 72 older children born in the first half of the year and 84 younger children born in second half of the year), found that the older group of L2 children performed better at receptive skills than the younger ones, albeit the age difference was minimal. The data of this quasi-experimental study were collected on three occasions. In the pretest phase, conducted at the end of year 1 (i.e., their first year of primary schooling), the participants’ receptive language was assessed using the British picture vocabulary scales, Macmillan’s individual reading analysis, sentence comprehension test, keywords reading list, and The Bus Story (to test oral comprehension). The posttest was conducted when the children completed year 2 and a follow-up test was administered when they commenced year 3. Statistical analysis showed that the older group of children performed consistently better than their younger counterparts in receptive vocabulary, reading accuracy, oral and reading comprehension. The researchers attributed this result to the older children’s “better developed cognitive skills” (p. 153), which allowed them to be more conscious of contextual cues.

According to Zdorenko and Paradis (2012), instructed L2 children constitute a unique group of learners because they learn the L2 via developmental acquisition rather than L1 transfer (i.e., they acquire the L2 by trial and error rather than transferring their knowledge from the L1). Rather than attributing the differences in learning rates between younger and older learners to linguistic factors, Dewaele, Petrides, and Furnham (2008) suggest that the difference occurs because of psychological and emotional factors, and specifically that those who start learning their second or even third or more language at a younger age (i.e., early AOA) have lower Foreign Language Anxiety (FLA) and a higher perceived level of oral proficiency. However, it should be noted that, unlike the studies by Muñoz (2006) and Sollars and Pumfrey (1999), this study concerns the participants’ self-perceived rather than measured proficiency.

Finally, other researchers explain that variation in ultimate attainment between adult L2 learners and child L2 learners occurs because of the distinct types of L2 knowledge each tends to develop. Ellis (2005b) explains that “[l]earners who began learning the L2 as a

Rhonda Oliver et al.

child are more likely to display high levels of implicit knowledge, whereas those who began as adolescents or adults—especially if they were reliant on instruction—are more likely to display high levels of explicit knowledge” (p. 152). Hence, future research investigating AOA may benefit from administering different tests designed to tap into different types of L2 knowledge because past research may have compared groups with different AOA on a type of knowledge that is L2-specific (i.e., explicit knowledge) rather than the type of knowledge native speakers possess (i.e., implicit knowledge) and that L2 instruction ultimately should aim at addressing.

Child Interaction

One line of research to emerge relatively early in Child ISLA was that based on the interaction hypothesis (Long, 1983). Long’s premise is that interaction “connects input, internal learner capacities, particularly selective attention, and output in productive ways” (1996, pp. 451–452). Adopting this interactionist approach, Child ISLA scholars investigated the pattern and pedagogical implications of negotiation for meaning and interactional feedback on the L2 child’s learning and did so by pairing learners with each other and with L1 peers, and by using a range of instructional tasks to prompt interaction (e.g., Mackey & Philp, 1998; Mackey & Silver, 2005; Mackey, Kaganas, & Oliver, 2007; Oliver, 1995; Oliver & Mackey, 2003; Pinter, 2007). A study by Oliver (1998) found that children do indeed negotiate for meaning, although using strategies in different proportions to adults. Comparing the negotiation by 196 L2 children (aged 8–13) with that by L2 adults reported in Long (1983), Oliver found that the child participants employed far fewer clarification requests (5.71% compared to 10.35%) and confirmation checks (5.72% compared to 18.15%). Furthering this work, and using the same data set, Oliver (2002) examined the effectiveness of pairing methods (32 nonnative speakers (NNS)—native speakers (NS), 48 NNS-NNS, and 16 NS-NS dyads) on the L2 children’s interactional patterns and amount of negotiation for meaning. The study found that NNS-NNS pairs tended to engage in the most negotiation for meaning, actively modifying their output to accommodate their conversational partners.

The positive impact of child interaction on L2 learning has been accounted for by the assistance gained from peers. In the context of Hungarian language education, where there is an emphasis on mechanical practice such as drilling, pattern practice, and expression memorization rather than on spontaneous and meaningful communication, Pinter (2007) conducted a small scale study in which two 10-year-old EFL Hungarian boys with low proficiency interacted with each other. The results showed that the learners accommodated each other’s communication needs and supported their partner by supplying unknown words, suggesting positive effects of peer interaction. Further, despite their hesitation and lack of fluency at first, over the course of the study both children reported feeling more confident and were better able to use communication strategies such as clarification requests to negotiate meaning. In an immersion context in Montreal, Canada, Ballinger (2015) investigated how Grade 3 and 4 learners interacted with each other when each other’s L1 was the respective target language (English or French). The analysis of 22.5 hours of interaction from eight pairs of learners suggested that the learners were able to reciprocally provide linguistic support for each other. Ballinger argued that, although social relationships between the child learners may have an impact on the ultimate benefit of collaborative interaction (e.g., feedback can be considered rude and thus its effectiveness may be lost), interaction between children with complementary language backgrounds is facilitative of L2 learning (see also Sato & Ballinger, 2016).

Related to research on child interaction are a number of Child ISLA task-based studies that have examined mediating variables. For example, Mackey et al. (2007) explored the impact of task familiarity in a study involving 40 ESL children aged 7–8 who had received English schooling in Australia for 10–14 months. The participants were put into pairs to perform communicative tasks whose content and procedural familiarity was controlled. The results indicated that the participants who were given unfamiliar tasks engaged in more negotiation for meaning than the dyads assigned familiar tasks—they asked more clarification requests, produced more confirmation checks, and corrected each other's non-target-like utterances more frequently. Also in Australia, Philp, Oliver, and Mackey (2006) explored the use of pretask planning and L2 children's learning outcomes. Forty-two ESL children from 5 to 12 years old from four Australian primary classrooms were given three communicative tasks over 3 weeks, with planning time ranging from zero, 2 minutes, and 5 minutes. The participants' fluency was measured by the number of false starts and reformulations while their accuracy was coded based on target-like communication units, and their complexity was assessed by the amount of subordination and coordination (grammatical complexity) and the number of lexical words (lexical complexity). Data analysis showed that pretask planning had little benefit as the children focused only on reciting their rehearsed utterances and appeared less interested in their partner's production of language. However, when no or little planning time was given, the participants provided more corrective feedback and modelled target-like output to each other, produced more words per minute, and negotiated the task using more complex language.

Attention in Child ISLA

One of the commonalities between Child (I)SLA and adult (I)SLA is the way in which attention is used to explain the learning process and to facilitate L2 development in the classroom. Due to the apparent similarities between Child SLA and L1 acquisition, some researchers in the 1980s and 1990s (e.g., Krashen, 1984; Truscott, 1996) questioned the need for L2 instruction especially for children. However, to date, research clearly indicates that both implicit and explicit techniques that draw learners' attention to language forms (e.g., form-focused instruction: FFI)—is beneficial and in many cases necessary for sustainable and accurate Child SLA. Harley's (1989) study is one of the earliest that investigated whether the teaching of grammar to children was beneficial. The study involved fifth- and six-graders (aged 10–12) in a French immersion context and the focused-input instruction concerned two French grammar points, the *imparfait* and *passé composé*. After 8 weeks, the immediate posttest showed that the experimental group (FFI) outperformed the control group (without FFI). However, there was no significant difference between the two groups in the delayed posttest conducted 3 months later. The results led Harley to conclude that FFI was useful to raise child learners' meta-linguistic awareness. However, its long-term effect remained doubtful. Day and Shapson (1991), in another key study, demonstrated the effect of attention to form on Child SLA. Also examining FFI, the researchers conducted an experiment with 12 French immersion classes of year 7 students (with a total of 315 students aged 12–13) from four districts in Vancouver, Canada. Over a period of 5–7 weeks (average of 17.4 hours of instruction), the experimental group was given treatment focusing on the conditional. The students were assisted in practising the grammatical structure with linguistic games and exercises that boosted their accuracy in communicative and formal, structured situations, for instance, discussion of futuristic elements and hypothetical examples. The control group also included immersion students, but they engaged simply in normal classroom instruction.

Rhonda Oliver et al.

The pretests and posttests comprised a cloze test of conditional forms, a composition about a comic character or a famous person they would like to be, and an oral interview about hypothetical situations. The results indicated that, in the posttest, consisting of the cloze test and composition, the experimental group outperformed the control group and continued to maintain higher gains in the follow-up test.

Since this early observation research, more contemporary investigations of FFI have evolved to include (quasi)experimental studies and these provide support for the effectiveness of FFI in different learning contexts. Tedick and Young (2014), for instance, investigated FFI within the context of immersion programs in the US where the educational focus is primarily on content rather than on linguistic forms and where, as a result of this approach, many Spanish-as-a-second-language children in the US tend to speak a “grammatically inaccurate” (p. 2) form of the language. Conducting a research project on the grammatical instruction of two Spanish past tenses, the *imperfect* and *preterit*, to fifth-graders (aged 10–11), Tedick and Young exposed the participants to seven lessons (approximately 6.5 hours) of FFI activities such as highlighting the two tenses in different colours in a biography, listing the verbs on a chart, and discussing the patterns. Data sources included quantitative data collected from pre-, post-, and delayed posttests with 10 focal students and qualitative data collected from classroom observations, field notes, and two teacher interviews. The results showed that the learners demonstrated development in their metalinguistic awareness and produced more target language tokens in the post-FFI observations.

In Southeast Asia, Shak and Gardner (2008) also studied the pedagogical benefits of FFI with 78 ESL children (aged 9–12) in Brunei Darussalam. In contrast to Tedick and Young, they extended FFI to include a rich variety of consciousness-raising activities including dictogloss (to elicit the output of *did + not + base form*) and other communicative pair- and group-work activities in a 2-day workshop. Different from other FFI projects, this study did not employ an experimental design; rather it assessed the participants’ perspectives with regard to FFI task enjoyment, ease, performance and motivation through an attitude questionnaire and group interviews. The findings showed that not only were the FFI tasks perceived by the children to be cognitively stimulating and enjoyable for language development, but the tasks were felt to have had a positive impact on their L2 learning. 83% of the children (after day 1) and 95% (after day 2) provided affirmative responses, reporting that they knew more about the form and functions of the target language structure and vocabulary as well as felt more confident about listening for information and sharing with friends through teamwork.

Another FFI technique that has received considerable attention in Child ISLA is corrective feedback. Like adult SLA research (see Dabaghi, 2011; de Vries, Cucchiari, Strik, & van Hout, 2011; Ellis, 2005a, 2011; Lee, 2013; Li, 2014; Lyster, Saito, & Sato, 2013; Mackey & Philp, 1998; Pawlak, 2013, 2014a, 2014b; Pawlak & Tomczyk, 2014; Qiao, 2013; Rassaei, 2013; Sheen, 2004, 2007; Shintani & Ellis, 2013), corrective feedback research began with observations of interaction between two or more children. Oliver (1995), for example, found that children, like adults, can provide their peers with feedback in the form of recasts (i.e., reformulation of non-target-like form to target-like, while maintaining the meaning) and that children use this feedback in their subsequent production. Specifically, 61% of learner error turns received feedback from NS child peers, with just over one-third comprising recasts and two-thirds negotiation for meaning; however, only approximately 10% of the recasts were incorporated in NNS subsequent responses. Nonetheless, a close examination of the data showed that it was either not possible (16% of the recasts) or not appropriate (55% of the

recasts) for the learners to do so. If these interactions were excluded, then the learners were found to use more than one-third of all the recasts provided.

Age effects have also been reported for the patterns of feedback provided to and then used by learners (i.e., uptake or modified output), characterizing the nature of child interaction and potential learning outcomes. Oliver (2000) compared ESL adults' and children's reactions to negative feedback. Using a task design that involved 32 NS-NNS dyads, the study found that the patterns of interaction were affected by age differences, and specifically, the adult NS interlocutors provided more implicit feedback in the form of recasts to their conversational partners than did the child NS interlocutors. Further, in reaction to feedback, the older learners were found to be better able to modify their non-target-like utterances. Adult ESL learners responded to negative feedback more frequently than the ESL child learners in both teacher-fronted lessons (29.1% and 21.1% respectively) and pair-work activities (32.8% and 24% respectively).

In a later study Oliver and Grote (2010) investigated recasts—specifically those that were multimove and single-move in three interactional contexts: teacher-ESL child learners, child NS-NNS students, and child NNS-NNS students (aged 7–13). Comparing the results of this research with Sheen's (2006) study, which “focused” on adults, Oliver and Grote reported that L2 children tended to provide and receive fewer multiple move recasts, but more single move recasts than adults. Further, the child learners had a lower level of uptake than did adult learners for all types of recasts in all three contexts. In Hungary, Pinter (2006) explored the way 10-year-old EFL children and college students performed information-gap tasks, namely spot the differences. She found that when children interacted in their L1, they spotted significantly more differences between the pictures than in L2. In contrast, the adult participants were more consistent in their interactions. This difference was explained by the fact that the adults produced more language and engaged in more checking, repeating each other's descriptions, asking for clarifications, co-constructing utterances, and negotiating misunderstandings. The children, on the other hand, tended to avoid difficult English words and employed significantly fewer negotiation strategies. In other words, the adult learners appeared more successful in L2 interaction than the child counterparts as far as feedback moves go.

Not only have child and adult L2 learners been found to differ, age effects have been reported for children of different ages. Oliver (2009) in a study of 32 younger children (aged 5–7 years) undertaking paired task work in class, found that although younger learners could negotiate and provide each other with feedback—much in the same way their older counterparts do—they were more concerned with “self” than “other” in their interactions, perhaps reflecting their egocentric stage of psychosocial development. Younger learners also were less bound by truthfulness and appeared to have a more flexible approach to undertaking the tasks. When Philp, Oliver, Philp, and Mackey (2008) compared the way ESL children aged 5–7 and 11–12 responded to on-task feedback, they found other age differences. For instance, they found that the older children, who had greater cognitive maturity, used teacher feedback to modify their output productions while, in contrast, the younger children were unable to modify their output unless scaffolded with pretask examples.

Another body of research related to the effectiveness of corrective feedback on Child SLA comes from classroom observational studies. In New Zealand, Choi and Li (2012) explored the uptake of corrective feedback by primary ESL children from years 2 to 6 during their class work. The study illustrated that although recasts and explicit correction were the teachers' preferred feedback methods, it was elicitation and clarification requests that resulted in optimum learner uptake (100%). While recasts received a

low uptake and repair rate for grammatical errors, they were a useful type of feedback for pronunciation errors. The analysis of the recorded lessons showed that the learners often indicated their uncertainty about the pronunciation of a word by pausing or raising their voice, signalling an expectation of feedback. Therefore, the teacher's recasts in this "form-focused dynamic" (p. 348) appeared to raise the students' awareness of the form and increase their noticing. In addition, phonological feedback led to the highest uptake (90%) and repair rate (87%), whereas lexical feedback the least (i.e., 64% and 44% respectively). In another classroom-based study, Oliver and Mackey (2003) investigated the type of feedback provided and the uptake of this feedback according to different lesson contexts. They conducted the study in five ESL classes with students aged 6–12 years. Over a period of 10 weeks, feedback was recorded in four contexts: (1) content (recasts comprised the dominant feedback type in this context), (2) classroom management (again recasts predominated), (3) communication (recasts and negotiation questions), and (4) explicit language instruction (metalinguistic commentary and explicit feedback). Data analysis revealed that feedback provided in the language-focused component of lessons resulted in 85% learner-modified output. Feedback in communicative contexts led to 38% modified output. Feedback in content-focused lessons resulted in only 27% modification of non-target-like forms while feedback provided in management contexts did not result in any uptake.

Other research has provided useful information to teachers about the various types of feedback, in particular that their preferences do not always align with learner needs, feedback effectiveness, or learner uptake. For example, Choi and Li (2012) found that although grammatical errors occurred the most frequently (157 errors), only approximately 50% received feedback. In contrast, phonological and lexical errors happened considerably less often (i.e., 48 and 27 respectively), but received a remarkably higher rate of teacher feedback (i.e., 81% and 93%). With regard to feedback types, Lyster and Mori (2006) reported that elementary teachers in both French and Japanese immersion relied predominantly on recasts (54–65%), while prompts and explicit correction were used to a significantly lesser extent (i.e., 26–38% and 7–9% respectively). These findings echoed those of Lyster and Ranta (1997) in their study of L2 children in Grades 4, 5, and 6. They found that French immersion teachers demonstrated a strong tendency to use recasts (accounting for more than 55% of the feedback provided). However, they also found that recasts comprised the lowest rate of learner-generated repair (only 31%). More effective types of corrective feedback included repetition of errors, metalinguistic feedback and clarification requests, which led to 78%, 86%, and 88% of uptake, respectively. Yet they were employed minimally by the teachers (i.e., 5%, 8%, and 11%).

The effectiveness of corrective feedback on Child SLA has also been examined using experimental designs with immediate posttests. Inspired by Harley (1989, 1998), Lyster (2004a, 2004b) conducted a series of quasi-experimental studies to explore the effectiveness of FFI (with or without feedback) in French classrooms in Canada. In the first, Lyster (2004a) examined the impact of FFI on French noun-endings and grammatical gender in a quasi-experimental project that involved 179 L2 fifth graders aged 10–11. Over an instructional period of 5 weeks, the study reported that FFI was the most effective when implemented in combination with prompts. In his next study (Lyster, 2004b), the lesson focus was perfect and imperfect past tenses. The children in this study were aged from 7 to 14 years. The posttests found that FFI was slightly less effective than recasts while prompts resulted in the most significant improvements, with the participants demonstrating the ability to self-repair without explicit provision of the target form.

The long term impact of interactional feedback on language acquisition has also been investigated in a series of studies employing pretest/posttest design. Mackey and Oliver (2002) undertook their study with 22 child ESL learners aged 8–12 from an intensive English centre in Australia. The study required the child participants to engage in interactional tasks with adult native speakers over 5 weeks with a focus on English question forms. The study found that 8 out of 11 children receiving feedback from the adult interlocutors demonstrated sustained development in the posttests whereas children receiving no or little feedback showed much slower progress in terms of question development. Based on Mackey and Oliver (2002), Mackey and Silver (2005) assigned 26 migrant children aged 6–9 in Singapore to two groups: the experimental group received interactional feedback during the pedagogical tasks that they performed with adult native speakers while the control group did not. The results showed significant statistical differences between the two groups: The children in the experimental group appeared to demonstrate substantial gains, exhibiting a marked improvement of their questioning skills.

In sum, the vibrant research of Child ISLA reveals that the process and product of Child SLA shares many similarities to, but at the same time is different from that of L1 acquisition and also from that of adult SLA. In addition, research indicates that child interaction differs in a number of ways not only from that of adult interaction, but even between children of different ages. These results may explain the differential effects of L2 instruction depending on the learners' age. The findings from Child ISLA studies suggest the importance of nuanced pedagogical intervention due to the fact that children are developing cognitively and socially in addition to acquiring L2 knowledge. In the following section, we will make some pedagogical suggestions based on these research findings.

Pedagogical Implications

Child ISLA classroom-based research has important pedagogical implications, informing L2 child instructional approaches in a number of ways including the way in which input is provided, the use of output activities, the organization of pair work, the effective provision of feedback, the importance of FFI, and the need to carefully consider individual differences.

First, research suggests the importance of different ways to maximize interaction through various designs and implementation of instructional tasks. Mackey et al. (2007) reported that both familiar and unfamiliar tasks (with regard to content and procedure) have their own benefits, the former eliciting more feedback and resulting in more output modification and the latter generating more clarification and confirmation and leading to more adjustments of linguistic forms. Therefore, Mackey et al. (2007) suggest that teachers should take account of task familiarity when planning interactional activities. This recommendation is consistent with Pinter (2007), who advises teachers to adopt task repetition to enhance students' confidence and fluency in the target language. Providing more general advice, Nicholas and Lightbown (2008) remind teachers that it is of paramount importance to take the child's social, cognitive, and linguistic differences into account when planning instruction. Each child not only has unique physiological and constitutional characteristics but also grows up in a different social and familial environment, which impacts on the expansion of their knowledge of the world and the developmental rate of their language to express new concepts.

Other researchers suggest it is advisable to train children to develop necessary task-related strategies such as clarification questions, repetition, and alternative ways of expressing meaning (Ballinger, 2015; Pinter, 2006, 2007). Sato and Ballinger (2012) trained Grade

3 and 4 learners to provide corrective feedback to each other and found that not only did the frequency of feedback increase over time, but the learners' language development was positively affected by the training.

Informed by a number of interactional studies, it is advised that teachers carefully consider the way they group their students. First, teachers need to take into account that the results may well be different for children than they are for adults: teachers "must be aware that the way that children are paired according to NNS-NS status and proficiency level is likely to influence the frequency with which negotiation occurs and that the pattern of interaction will not necessarily be the same as it is for adults" (Oliver, 2002, p. 108). Second, the organisation of pairs and groups can be done in ways that enhance learning. For example, Van den Branden (2008) reported that putting L2 children into small groups can significantly reduce social threat and facilitate interaction. Other advice is provided by Pinter (2007), who suggests pairing students with the same partner for an extended period of time to enhance their level of comfort and confidence. In short, language teachers need to not only provide extensive input but also optimize opportunities for L2 output and interaction through effective organization of collaborative activities (see Philp & Duchesne, 2016; Sato, 2016).

While interaction and production tasks obviously play a vital role in Child SLA, the results from FFI studies suggest that L2 child pedagogy should not overlook attention-directing activities that promote learners' language awareness, which in turn facilitates the development of linguistic accuracy. Although some children may achieve a high competence in the target language in natural settings, suggesting that FFI is unnecessary in Child SLA, most children learning a second or foreign language in a classroom context cannot achieve such linguistic success on their own, hence the pivotal role of FFI (Spada & Lightbown, 2008). In this sense, Harley (1993) argues that a balance needs to be struck between FFI and meaningful communication because an overconcentration on accuracy and the overuse of an analytic approach will inhibit L2 children from producing the target language, limit their risk-taking willingness, and negatively influence the development of the confidence necessary for successful L2 communication. As Lyster (2007) argues, balancing the focus on language and communication depending on types of learners is the key to successfully supporting their attentional shift (see also Sato, 2011).

More specific FFI activities suggested in the Child ISLA literature include the use of derivational morphology through reading aloud from story books (Lyster, 2015), consciousness-raising tasks, dictogloss, grammar interpretation, and grammaring (Shak & Gardner, 2008). Ellis and Shintani (2013) suggest that FFI can include not only rule-based grammar, but also instruction concerned with formulaic expressions. As reported by both Harley (1989) and Day and Shapson (1991), the teachers in their studies tended to focus on the exciting and intrinsically motivating communicative activities at the expense of FFI, resulting in the production of simplified forms and an overreliance on communicative strategies. According to Lyster (2004b), this practice can potentially limit the L2 child's interlanguage development. Lyster goes on to suggest that in linguistically difficult areas where persistent errors are made, teachers may need to include more than just meaning-focused activities, and that explicit instruction of the target forms may also be required.

Research on the effect of age-related differences foregrounds the importance of varying feedback, instructional methods and pedagogical tasks when teaching children of different ages (Oliver et al., 2008; Philp, Oliver, & Mackey, 2008; Pinter, 2006). With regard to teaching younger ESL children (aged 5–7), Oliver et al. (2008) suggest using pretask examples as a way of whole-class scaffolding rather than providing on-task feedback. In contrast, older ESL children (aged 11–12) may benefit more from the latter method as they are mature

enough to notice interaction-based feedback (Oliver et al., 2008). Oliver (1998) and Pinter (2006) assert that tasks used with one age group need to be modified to suit the psycho-developmental characteristics and interests of another.

Teaching Tips

- Remember that Child SLA happens concurrently with the child's cognitive and psychological development.
- Grammatical intervention, including explicit types, does facilitate Child SLA because L2 and L1 acquisition are different.
- Type of tasks influences the way children interact with the teacher and with each other; therefore, it is important to choose tasks that are both appropriate and timely.
- Don't forget that learners of different ages benefit from the same instruction in different ways.
- Training learners how to interact with each other in more effective ways may be beneficial for maximizing the effectiveness of interaction in L2 learning.
- Learners' attention should be directed in ways to achieve higher grammatical competency; however, there needs to be a balance to allow learners to focus on meaningful communication.

Future Directions

Research on Child ISLA is distinct from that conducted with other groups of participants. As discussed at the beginning of the chapter, researching Child ISLA requires particular efforts because of various theoretical and methodological issues. In this vein, Pinter and Zandian (2014) provide a list of factors that Child ISLA researchers need to take into consideration, including understanding children's perspectives, obtaining consent, establishing a friendly relationship, and optimizing learner output. Overall, there is a need for researchers to abide by the principles of reducing power distance and building trust, confidence, and rapport (Pinter, Kuchah, & Smith, 2013). In particular, Child ISLA researchers may want to involve children as "active participants" or "co-researchers," thereby allowing them to discuss issues of their concern to be investigated (Pinter, 2014; Pinter et al., 2013). As with L2 teaching pedagogy where student-centred approaches are valued and prioritized, Pinter et al. (2013) argue that Child ISLA researchers should move toward child-centred approaches. In this regard, Pinter and Zandian (2014) argue that because "children are 'experts' of their own lives" (p. 64), alternative ways have to be developed to research "with" children rather than "on" children. However, researchers contemplating this child-focused approach have to take into account the fact that a "genuine barrier to children engaging in research is their lack of research knowledge and skills" (Kellett, 2010, p. 197), as well as their level of cognitive and social maturity. Therefore, children need to be trained if they are to be placed at the centre of the research process.

Along with ethical, effective, and rigorous methodology, much more research is needed to understand both the similarities and differences in the trajectory and nature of differently aged learners particularly in instructional settings. One line of research that allows for the close intersection between theory and practice is in the areas of FFI and classroom interaction. To date most studies have employed short-term observational designs. However,

Rhonda Oliver et al.

long-term, (quasi)experimental research studies are also needed to explore whether FFI definitively results in increased accuracy (Tedick & Young, 2014). When conducting a classroom-based study, however, one needs to consider multiple factors apart from the instructional variables alone that may influence L2 children's linguistic progress. As Lyster (2004b) points out, the effects of form-focused interventions can be neutralized due to the linguistic nature of the language item (i.e., some items are more complex and thus harder to acquire than others). Therefore, when undertaking FFI research with children, there is a need to carefully control variables and to interrogate in depth the interactional data that emerges during the data collection process. Like adult SLA research, there is also a need to develop rigorous ways to accurately measure L2 children's acquisition. At present, questions still remain as to whether acquisition is demonstrated merely by learner uptake and production (and if so, how many instances of the targeted structure should occur), and whether posttesting learners' identification or use of targeted structures is appropriate in the Child SLA context.

Finally, future studies need to explore the vexed question of form versus content in language classrooms. Is it possible, for instance, to have an effective cross-curricular pedagogy that supports L2 children's acquisition of linguistic forms while teaching content? As Lightbown (2014) asserts, the separation of language and content "may deprive students of opportunities to focus on specific features of language at the very moment when their motivation to learn them may be at its highest" (p. 30). How this intersection of form and content teaching might be achieved in various contexts is another area worthy of future research. For example, it might be fruitful for researchers to investigate child learners' spontaneous production of forms in content-oriented activities (Tedick & Young, 2014). It would also be useful to examine aspects of interaction that can mediate the potential for language change (Philp & Duchesne, 2008). Clearly there are numerous areas of Child ISLA that would benefit from further research. Given the paucity of research in this area compared to that work conducted with other aged groups, there is much work to be done. Although it is a challenging area, it is also one offering considerable rewards.

References

- Ballinger, S. (2015). Linking content, linking students: A cross-linguistic pedagogical intervention. In J. Cenoz & D. Gorter (Eds.), *Multilingual education: Between language learning and translanguaging* (pp. 35–60). Cambridge: Cambridge University Press.
- Chilla, S., Habertzettl, S., & Wulff, N. (2013). Dummy verbs in first and second language acquisition in German. In E. Blom, I. van de Craats, & J. Verhagen (Eds.), *Dummy auxiliaries in first and second language acquisition* (pp. 209–250). Berlin/Boston, MA: De Gruyter Mouton.
- Choi, S.-Y., & Li, S. (2012). Corrective feedback and learner uptake in a child ESOL classroom. *RELC Journal*, 43(3), 331–351.
- Cook, V. (2010). The relationship between first and second language learning revisited. In E. Macaro (Ed.), *The continuum companion to second language acquisition* (pp. 137–157). New York: Continuum.
- Cornips, L. (2013). Child use of auxiliary + infinitive in Dutch: Acquisition device or reflection of the input. In E. Blom, I. van de Craats, & J. Verhagen (Eds.), *Dummy auxiliaries in first and second language acquisition* (pp. 369–394). Berlin/Boston, MA: De Gruyter Mouton.
- Dabaghi, A. (2011). Second language acquisition and the role of corrective evidence. *Language and Literature: European Landmarks of Identity*, 8, 22.
- Day, E., & Shapson, S. (1991). Integrating formal and functional approaches to language teaching in French immersion: An experimental study. *Language Learning*, 41, 25–58.

- de Vries, B. P., Cucchiaroni, C., Strik, H., & van Hout, R. (2011). Adaptive corrective feedback in second language learning. In S. De Wannemacker, G. Clarebout, & P. De Causmaecker (Eds.), *Interdisciplinary approaches to adaptive learning: A look at the neighbours* (pp. 1–14). Berlin: Springer.
- Dewaele, J.-M., Petrides, K. V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning*, 58(4), 911–960.
- Dixon, L. Q., Zhao, J., Shin, J. Y., Wu, S., Su, J. H., Burgess-Brigham, R., Gezer, M. U., & Snow, C. (2012). What we know about second language acquisition: A synthesis from four perspectives. *Review of Educational Research*, 82(1), 5–60.
- Dulay, H., & Burt, M. (1974). A new perspective on the creative construction process in child second language acquisition. *Language Learning*, 24(2), 253–278.
- Ellis, R. (2005a). *Instructed second language acquisition: A literature review*. Wellington: Ministry of Education.
- Ellis, R. (2005b). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172.
- Ellis, R. (2011). Epilogue: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, 32, 335–349.
- Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language acquisition research*. Abingdon: Routledge.
- Foster-Cohen, S. (2010). Second language acquisition and the younger learner: Child's play? Book review. *Studies in Second Language Acquisition*, 32(1), 153–154.
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.
- Geva, E., & Yaghoub Zadeh, Z. (2006). Reading efficiency in native English-speaking and English-as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading*, 10, 31–57.
- Harley, B. (1989). Functional grammar in French immersion: A classroom experiment. *Applied Linguistics*, 10, 331–359.
- Harley, B. (1993). Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition*, 15, 245–259.
- Harley, B. (1998). The role of form-focused tasks in promoting child L2 acquisition. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 156–174). Cambridge: Cambridge University Press.
- Haznedar, B., & Gavruseva, E. (2008). Recent perspectives in child second language acquisition. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 3–16). Amsterdam: John Benjamins.
- Huang, J., & Hatch, E. M. (1978). A Chinese child's acquisition of English. In E. M. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 76–88). Rowley, MA: Newbury House.
- Ionin, T. (2008). Progressive aspect in child L2-English. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 17–54). Amsterdam: John Benjamins.
- Iwasaki, J. (2008). Acquiring Japanese as a second language (JSL) in a naturalistic context. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 231–254). Amsterdam/Philadelphia: John Benjamins.
- Kellett, M. (2010). Small shoes, big steps! Empowering children as active researchers. *American Journal of Community Psychology*, 46(1–2), 195–203.
- Krashen, S. D. (1984). Immersion: Why it works and what it has taught us. *Language and Society (special issue)*, 12, 61–64.
- Lakshmanan, U. (1994). *Universal grammar in child second language acquisition: Null subjects and morphological uniformity*. Amsterdam: John Benjamins.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. London: Longman.

- Lee, E. J. (2013). Corrective feedback preferences and learner repair among advanced ESL students. *System*, 41, 217–230.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.
- Li, S. F. (2014). Oral corrective feedback. *ELT Journal*, 68(2), 196–198.
- Lightbown, P. M. (2014). *Focus on content-based language teaching*. Oxford: Oxford University Press.
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th ed.). Oxford: Oxford University Press.
- Long, M. H. (1983). Linguistic and conversational adjustments to non-native speakers. *Studies in Second Language Acquisition*, 5, 177–193.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12(3), 251–285.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2: Second language acquisition* (pp. 413–468). New York: Academic Press.
- Lyster, R. (2004a). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399–432.
- Lyster, R. (2004b). *How to predict the grammatical gender of 8000 French nouns*. Paper presented at the meeting of the American Association for Applied Linguistics, Portland, OR.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam: John Benjamins.
- Lyster, R. (2015). Using form-focused tasks to integrate language across the immersion curriculum. *System*, 54, 4–13.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28(2), 269–300.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classroom. *Studies in Second Language Acquisition*, 19, 37–66.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching: Surveys and Studies*, 46(2), 1–40.
- Mackey, A., Kaganas, A. P., & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly*, 41(2), 285–312.
- Mackey, A., & Oliver, R. (2002). Interactional feedback and children's L2 development. *System*, 30(4), 459–477.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *Modern Language Journal*, 82(3), 338–356.
- Mackey, A., & Silver, R. E. (2005). Interactional tasks and English L2 learning by immigrant children in Singapore. *System*, 33, 239–260.
- Montrul, S. (2004). *The acquisition of Spanish: Morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition*. Philadelphia, PA: John Benjamins.
- Montrul, S. (2006). Universal grammar in second language acquisition: A history. *Studies in Second Language Acquisition*, 28(4), 642–643.
- Muñoz, C. (2006). The effects of age on foreign language learning: The BAF project. In C. Muñoz (Ed.), *Age and the rate of foreign language learning* (pp. 1–40). Clevedon: Multilingual Matters.
- Nicholas, H., & Lightbown, P. M. (2008). Defining child second language acquisition, defining roles for L2 instruction. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 27–52). Amsterdam/Philadelphia: John Benjamins.
- Oliver, R. (1995). Negative feedback in child NS/NNS conversation. *Studies in Second Language Acquisition*, 17, 459–483.
- Oliver, R. (1998). Negotiation of meaning in child interactions: The relationship between conversational interaction and second language acquisition. *Modern Language Journal*, 82(3), 372–386.
- Oliver, R. (2000). Age difference in negotiation and feedback in classroom and pair work. *Language Learning*, 50(1), 119–151.

- Oliver, R. (2002). The patterns of negotiation for meaning in child interactions. *Modern Language Journal*, 86(1), 97–111.
- Oliver, R. (2009). How young is too young? Investigating negotiation of meaning and corrective feedback in children aged five to seven years. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language interaction research in honour of Sue M. Gass* (pp. 135–156). New York: Routledge.
- Oliver, R., & Grote, E. (2010). The provision and uptake of different types of recasts in child and adult ESL learners: What is the role of age and context? *Australian Review of Applied Linguistics*, 33(3), 26.1–26.22.
- Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *Modern Language Journal*, 87(4), 519–533.
- Oliver, R., Philp, J., & Mackey, A. (2008). The impact of teacher input, guidance and feedback on ESL children's task-based interactions. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 131–147). Amsterdam: John Benjamins.
- Ortega, L. (2001). Current options in graduate-level introductory SLA textbooks. *Second Language Research*, 17(1), 71–89.
- Park, K. (2014). *Information structure and dative word-order alternations in English and Korean: L1 children, L2 children, and L2 adults* (ProQuest Dissertations and Theses). Ann Arbor, MI.
- Pawlak, M. (2013). The effect of explicit and implicit corrective feedback on eliminating pronunciation errors. In E. Wanek-Klimczak & L. R. Shockey (Eds.), *Teaching and researching English accents in native and non-native speakers, second language learning and teaching* (pp. 85–101). Berlin: Springer.
- Pawlak, M. (2014a). *Error correction in the foreign language classroom: Reconsidering the issues*. London: SpringerLink.
- Pawlak, M. (2014b). Investigating learner engagement with oral corrective feedback: Aims, methodology, outcomes. In A. Lyda & K. Szcześniak (Eds.), *Awareness in action, second language learning and teaching* (pp. 69–84). Switzerland: Springer International.
- Pawlak, M., & Tomczyk, E. (2014). Differential effects of input-providing and output-inducing corrective feedback on the acquisition of English passive voice. In E. Piechurska-Kuciel & E. Szymanska-Czaplak (Eds.), *Language in cognition and affect, second language learning and teaching* (pp. 133–149). Berlin: Springer.
- Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.
- Philp, J., & Duchesne, S. (2008). When the gate opens: The interaction between social and linguistic goals in child second language development. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 83–104). Amsterdam: John Benjamins.
- Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics*, 36, 50–72.
- Philp, J., Oliver, R., & Mackey, A. (2006). The impact of planning time on children's task-based interactions. *System*, 34(4), 547–565.
- Philp, J., Oliver, R., & Mackey, A. (2008). *Second language acquisition and the younger learner: Child's play?* Amsterdam/Philadelphia: John Benjamins Pub. Co.
- Pica, T. (2005). Second language acquisition research and applied linguistics. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 263–280). Mahwah, NJ: Lawrence Erlbaum.
- Pinter, A. (2006). Verbal evidence of task related strategies: Child versus adult interactions. *System*, 34(4), 615–630.
- Pinter, A. (2007). Some benefits of peer-peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research*, 11(2), 189–207.
- Pinter, A. (2014). Child participant roles in applied linguistics research. *Applied Linguistics*, 35(2), 168–183.

Rhonda Oliver et al.

- Pinter, A., Kuchah, K., & Smith, R. (2013). Researching with children. *ELT Journal*, 67(4), 484–487.
- Pinter, A., & Zandian, S. (2014). “I don’t ever want to leave this room”: Benefits of researching “with” children. *ELT Journal*, 68(1), 64–74.
- Pinter, A., & Zandian, S. (2015). “I thought it would be tiny little one phrase that we said, in a huge big pile of papers”: Children’s reflections on their involvement in participatory research. *Qualitative Research*, 15(2), 235–250.
- Qiao, Z. (2013). Effects of corrective feedback on the acquisition of Chinese directional verb compounds by second language learners. *US—China Foreign Language*, 11(1), 46–57.
- Rassaei, E. (2013). Corrective feedback, learners’ perceptions, and second language development. *System*, 41(2), 472–483.
- Rocca, S. (2007). *Child second language acquisition: A bi-directional study of English and Italian tense and aspect morphology*. Amsterdam: John Benjamins.
- Sato, M. (2011). Constitution of form-orientation: Contributions of context and explicit knowledge to learning from recasts. *Canadian Journal of Applied Linguistics*, 14(1), 1–28.
- Sato, M. (2016). Interaction mindsets, interactional behaviors, and L2 development: An affective-social-cognitive model. *Language Learning*. Advanced online publication. doi:10.1111/lang.12214
- Sato, M., & Ballinger, S. (2012). Raising language awareness in peer interaction: A cross-context, cross-method examination. *Language Awareness*, 21(1–2), 157–179.
- Sato, M., & Ballinger, S. (Eds.). (2016). *Peer interaction and second language learning: Pedagogical potential and research agenda*. Amsterdam: John Benjamins.
- Shak, J., & Gardner, S. (2008). Young learner perspectives on four focus-on-form tasks. *Language Teaching Research*, 12(3), 387–408.
- Sheen, Y. (2004). Corrective feedback and learners uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263–300.
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10(4), 361–392.
- Sheen, Y. (2007). The effects of corrective feedback, language aptitude and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301–322). Oxford: Oxford University Press.
- Shibata, K., Watanabe, T., Sasaki, Y., & Kawato, M. (2011). Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science*, 9, 1413–1415.
- Shintani, N., & Ellis, R. (2013). The comparative effect of direct written corrective feedback and meta-linguistic explanation on learners’ explicit and implicit knowledge of the English indefinite article. *Journal of Second Language Writing*, 22(3), 286–306.
- Simon, E. (2010). Second language acquisition and the younger learner: Child’s play? *Journal of Child Language*, 37(4), 945–951.
- Sollars, V., & Pumfrey, P.D. (1999). Reciprocal and transmission models of teaching in E2L with young learners. *International Journal of Early Years Education*, 7, 141–157.
- Spada, N. (2015). SLA research and L2 pedagogy: Misapplications and questions of relevance. *Language Teaching*, 48(1), 69–81.
- Spada, N., & Lightbown, P.M. (2008). Form-focused instruction: Isolated or integrated? *TESOL Quarterly*, 42(2), 181–207.
- Spada, N., & Lightbown, P.M. (2010). Instructed second language acquisition. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed., pp. 108–123). London: Edward Arnold.
- Spyrou, S. (2011). The limits of children’s voices: From authenticity to critical, reflexive representation. *Childhood: A Global Journal of Child Research*, 18(2), 151–165.
- Tedick, D.J., & Young, A.I. (2014). Fifth grade two-way immersion students’ responses to form-focused instruction. *Applied Linguistics*, 1–25. Advanced online publication. doi:10.1093/applin/amu066
- Truscott, J. (1996). The case against grammar correction in L2 classes. *Language Learning*, 46(2), 327–369.

- Unsworth, S. (2007). L1 and L2 acquisition between sentence and discourse: Comparing production and comprehension in child Dutch. *Lingua*, 117(11), 1930–1958.
- Van den Branden, K. (2008). Negotiation of meaning in the classroom: Does it enhance reading comprehension? In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 149–172). Amsterdam/Philadelphia: John Benjamins.
- Zdorenko, T., & Paradis, J. (2012). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, 32(1–2 [99]), 38–62.

Instructed Heritage Language Acquisition

Silvina Montrul and Melissa Bowles

Background

In the last decades in particular, the US, like many other countries in the world, has opened its doors to many immigrants who are speakers of other languages. This situation has resulted in increasing enrollments of heritage speakers (HSs) in classes that teach world languages as foreign languages (Carreira & Kagan, 2011). HSs are bilinguals whose native, heritage language is a minority language. Although the heritage language is spoken at home as in monolingual acquisition, very often this language does not develop to the same extent as the majority language HSs acquire as a second language. As a result, HSs' ultimate attainment of heritage language proficiency in early adulthood is highly variable, as is typical in second language acquisition (Montrul, 2008). When input is restricted and insufficient before puberty, the development of the heritage language (HL) is delayed and interrupted, and young adult HSs' grammars often display properties typical of developmental stages of first (L1) and second language (L2) acquisition (Montrul, 2016). By adolescence and young adulthood, the HL, which has become secondary, now also manifests many of the same characteristics of the interlanguage systems of L2 learners. Due to apparent linguistic similarities despite differences in language learning experience between HSs and L2 learners, in the last two decades the field of heritage language acquisition has made substantial strides in theoretical and empirical research aimed at identifying how L2 learners and HSs compare or differ in their linguistic competence and processing abilities (Au, Knightly, Jun, & Oh, 2002; Keating, VanPatten, & Jegerski, 2011; Montrul, 2010). Socioculturally oriented research has addressed how affiliation and degree of allegiance with the ethnolinguistic community guide HL development (He, 2006, 2010) as well as how HSs' attitudes to the heritage language and culture and their social networks enhance or hinder HL development (Jo, 2001).

Like L2 learners, HSs need motivation to maintain and develop their HL through language and literacy instruction beyond what they acquire at home. Many young adult HSs seek to improve their basic grammatical and communicative knowledge of the HL, and it is very common to find HL learners in L2 and foreign language classes in North American and European universities (Håkansson, 1995). While some countries offer "mother tongue" (i.e., "heritage language") instruction in public schools (e.g., in Scandinavia), in many universities HSs take

regular foreign language classes with L2 learners, creating a challenging situation for instructors. Some have argued that curricula designed for foreign language students are largely inappropriate for HL students (Carreira & Kagan, 2011; Oh & Au, 2005; Peyton, Carreira, Wang, & Wiley, 2008), and others have even shown that being in classrooms with L2 learners may cause HL learners to feel self-conscious about their language ability (Potowski, 2002) and could affect learning outcomes. Nevertheless, the *outcomes* of classroom HL instruction have been largely understudied. That is, although there are indications that some aspects of HL learners' language development may differ from that of L2 learners due to their different language learning experience, there is little empirical evidence on the specific ways in which they differ and on how those differences may affect learning outcomes in instructed contexts. Whereas the field of instructed second language acquisition (ISLA) has enjoyed decades of systematic research on learner-internal and learner-external factors that make instruction maximally effective for instructed L2 learners (Loewen, 2015), there has not yet been such systematic research on instructed heritage language acquisition (IHLA). Even basic questions, such as whether instructed HL learners make learning gains compared to uninstructed HL learners who merely receive naturalistic input in the language in their homes and communities, remain unanswered. Given the increasing numbers of learners enrolling in classes in their HL, Bowles (in press) has argued that IHLA should be a field in its own right, taking as a point of departure research paradigms in ISLA.

In this chapter, we review the scant existing research on IHLA, addressing the following questions: Does formal instruction contribute to the linguistic development of HSs? What aspects of instruction are beneficial for HL learning? And related to this general question, how does instruction affect HL and L2 learners' language development? Finally, is it beneficial for L2 learners and HL learners to share the same classroom? We conclude by arguing that in order to move the field forward and for language instruction to be maximally effective for HL learners, systematic research on a variety of language domains must be conducted. We also propose a research agenda for such systematic study and draw pedagogical implications from the current knowledge base.

Key Concepts

Heritage speaker: A person who is to some degree bilingual in a minority language (the heritage language) and the majority language.

Minority language: A nondominant language in a particular society, which typically has lower status and less prestige than the dominant societal language. In the US, Korean, Russian, and Vietnamese are all minority languages, even though they are majority languages in Korea, Russia, and Vietnam, respectively.

Majority language: The dominant language in a society, which often has the status of an official language. In the US, English is the majority language.

Explicit knowledge: Knowledge "about" language that learners can verbalize, either with or without grammatical terminology. It is conscious and accessing it is slow.

Implicit knowledge: Unconscious knowledge "of" language that learners are unable to verbalize. It is retrieved very quickly and can be used for automatic processing.

Focus on form: An approach to instruction that consists of learners' attention briefly being drawn to some aspect of language form in a larger meaning-focused context.

Language-related episodes (LREs): Instances during meaning-based interaction when learners spontaneously focus on linguistic form.

Empirical Evidence

Does Instruction Contribute to HL Development?

Instruction aims to facilitate language learning by maximizing the role of input and output in the classroom. An important difference between HL acquisition and L2 acquisition is that the former usually happens in a naturalistic setting in early childhood, whereas the latter primarily takes place in the classroom later in childhood or adolescence. As with any early child language acquisition context, in HL acquisition there is typically little to no explicit instruction or information about grammaticality, in contrast to the classroom L2 acquisition setting. A central question in ISLA acquisition is what types of linguistic input and learning environments are most beneficial for L2 learners, and whether explicit instruction helps learners develop and restructure their linguistic systems. Many researchers argue that negative evidence—information regarding the impossibility of certain linguistic structures in the language being acquired—is not necessary and perhaps not even consistently available for bilingual and L1 acquisition (Pinker, 1989). However, the research on L2 acquisition that started in immersion contexts suggested that positive evidence alone may not be sufficient for the acquisition of certain L1-L2 contrasts or structures that are not present in the L1 or for the unlearning of developmental errors (Lightbown, 1998; Long, 1996; Trahey & White, 1993; White, 1991). That is, to draw their attention to particular linguistic features that otherwise might go unnoticed, L2 learners may benefit from occasional focus on form in the context of meaning-based communication. Several meta-analyses have additionally found that more explicit approaches to instruction, such as those that include explicit grammatical explanation or rule presentation, can be more beneficial for aspects of morphosyntax than implicit approaches (Norris & Ortega, 2000; Russell & Spada, 2006; Spada & Tomita, 2010).

An important general question is whether HL instruction promotes HL development and maintenance more than naturalistic language exposure in the home and/or community alone. Bylund and Díaz (2012) offer an answer to this question. In Swedish public schools, where this study was conducted, HSs receive heritage language instruction, which in Scandinavia is referred to as *mother tongue instruction*, at least one hour per week. The authors investigated the effects of weekly instruction on overall HL proficiency in two groups of twelfth grade (high school) Spanish HSs. One group ($n = 28$) was receiving HL instruction twice a week (2 hours total). The other group ($n = 26$) had received HL instruction until the eleventh grade, but for scheduling conflict reasons was no longer attending HL classes but was instead being instructed only in Swedish. The authors controlled for many variables, including participants' age of arrival in Sweden, length of residence in Sweden, amount of use of Swedish and Spanish, and amount of previous HL instruction, the only difference being that the uninstructed group had not received HL instruction for 10 months. Students who were receiving HL instruction at the time of testing outperformed students not taking HL classes that year on a written grammaticality judgment test and a written cloze test. Bylund and Díaz interpret these results to suggest that continued HL instruction through literacy development contributes to language maintenance and prevents L1 attrition at a critical time for language development (before the critical period), following Bylund's (2009) and Montrul's (2008) hypothesis of age effects in L1 loss. However, because the two groups were not tested a year earlier, it is impossible to know whether they were of similar proficiency in the eleventh grade. Furthermore, as Bylund and Díaz acknowledge, because the two measures tested written proficiency and metalinguistic ability to some extent, it is not clear whether

the HL learners who were continuing with classes had restructured their implicit system or whether they had simply gained explicit knowledge. Measures tapping implicit knowledge, such as oral production, would have to be used to corroborate this possibility. Furthermore, delayed posttests would be needed to determine whether the knowledge acquired through classroom instruction is retained over time, because the ultimate goal of instructed heritage or second language acquisition is long-term, rather than immediate, gains. Despite leaving these questions unanswered, Bylund and Diaz show that HL instruction contributes to HL morphosyntactic development, compared to no instruction.

If HL instruction in general makes a difference, the next question is whether the linguistic knowledge that HL learners bring to the classroom gives them an advantage in the classroom, as compared to L2 learners who do not bring such knowledge. A more nuanced question relevant for some HLs (e.g., Arabic and Greek) is whether the colloquial varieties to which HSs are exposed at home help them in learning the standard varieties of the language imparted in most classrooms. Albirini (2014) sought to address this question with HSs who spoke colloquial Palestinian or Egyptian Arabic at home and were receiving HL instruction in Modern Standard Arabic (MSA). Crucially, the colloquial varieties of Arabic differ in substantial ways both from MSA and from each other at the phonological, morphological, syntactic, and lexical levels. Arabic-speaking children learn a colloquial variety from birth and begin exposure to and acquisition of MSA when they enter school. Therefore, Arabic HSs typically know a colloquial variety of Arabic and although they may have heard some MSA on TV, because it is used in news and other media broadcasts on central satellite networks such as Al Jazeera and Al Arabiya, they do not typically have a command of it.

Albirini (2014) tested instructed Arabic HL and L2 learners' knowledge of sentential negation in MSA to determine whether knowledge of a colloquial variety provides an initial advantage for HL learners over L2 learners and whether that advantage is sustained as proficiency increases. (Although sentential negation differs in the colloquial varieties and in MSA, there are overall similarities between the colloquial and standard systems). Albirini tested 19 HL learners and 10 L2 learners in elementary MSA classes and 16 HL learners and 18 L2 learners in advanced MSA classes. All participants completed five oral tasks targeting negation in various contexts.

Results showed that in the elementary class, HSs had an advantage over L2 learners of Arabic, because their sentences involving negation were for the most part syntactically well-formed in MSA compared to the L2 learners' sentences. At least 60% of the low proficiency HL learners' errors could be attributed to transfer from their colloquial variety. However, the initial advantage appears to dissipate as HSs advance in proficiency, because among the students in the advanced class there was no significant advantage for HSs compared to L2 learners with sentential negation.

In conclusion, HSs' linguistic knowledge does appear to confer some advantages (at least initially) when they come to the HL classroom compared to L2 learners. Instruction also seems to contribute to HL development, although further research is needed, with assessments tapping both explicit and implicit knowledge, administered both immediately after instruction and some time later, to develop a fuller picture of the nature of that development.

What Aspects of Instruction Are Beneficial for HL Learning?

Linguistically oriented research has uncovered many grammatical areas that are underdeveloped in HL grammars, such as inflectional morphology (Montrul, 2016). Once researchers identify such areas that are hard to master and even stabilize at nontarget levels, the next

question is whether focused instruction furthers HL learners' development in those linguistic areas.

Among the first to address the role of explicit instruction in HL acquisition were Song, O'Grady, Cho, and Lee (1997), who showed that child HSs of Korean had difficulty using case markers to comprehend agent-patient relationships in Korean SOV and OSV sentences. They designed a 2-week intervention consisting of explicit explanations and examples of when to use the relevant case markers in Korean SOV and OVS sentences, and children took a posttest immediately after the instruction and a delayed posttest 9 weeks later. Results showed that children improved by more than 100% in accurately recognizing who was doing what in sentences with OSV order, from 25% accuracy on the pretest to 66.3% accuracy on the immediate posttest. This knowledge was mostly retained at the time of the delayed posttest, when accuracy was 56%.

Montrul and Bowles (2010) is another study that addressed whether focused instruction on specific grammatical targets helps HSs advance in their grammatical development of that specific area. They investigated the effects of instruction on college-age HL learners' knowledge of differential object marking (the preposition "*a*" in *Juan vio a María* "John saw María") and dative case marking in Spanish (the preposition "*a*") with psychological verbs like *gustar* "like" (*A Juan le gusta el fútbol* "Juan likes soccer"), which are also problematic for L2 learners (Bowles & Montrul, 2009; Guijarro Fuentes, 2012). Forty-five Spanish HL learners completed written production and acceptability judgment tests both before and immediately after instruction, which began with an explicit grammatical explanation of the targeted structures. After reading the grammatical explanation, learners completed a 20-item practice exercise online for each construction (*a*-personal, indirect objects, and dative experiencers with *gustar*-type verbs). Each practice item consisted of a sentence with a drop-down menu immediately preceding the object, from which the learners chose either *a* or a space, the latter indicating the absence of *a*. Participants received immediate, explicit feedback after each selection that indicated whether their response was correct and provided an explanation.

Results showed that instructed, but not uninstructed HL learners, made significant pre-posttest gains on the production test on all sentence types. However, instructed HL learners' gains were not equal in all areas or on all sentence types. Most notably, instruction did not affect their acceptability ratings as much as their production. On most sentence types, instruction affected acceptability ratings in the expected direction. However, on ungrammatical sentences with animate objects (missing the *a*-personal), as in (1), instructed HL learners' pre- and post-instruction acceptability ratings were not significantly different from each other.

- (1) *Pedro conoce el chef.
Pedro knows the chef.

Overall, Montrul and Bowles found that explicit instruction and feedback was highly beneficial to HL learners. In fact, the magnitude of the gains on all of the structures was higher for HL learners than for the L2 learners in Bowles and Montrul (2009), which followed the same design. Taken together, the results of the two studies suggest that negative evidence plays a role in both instructed L2 acquisition and HL acquisition, and that explicit instruction is beneficial for both groups, although the studies' design does not allow us to determine the individual contributions of explicit grammatical information and negative evidence.

Two other studies, Potowski, Jegerski, and Morgan-Short (2009) and Torres (2013), have gone a step further, investigating how HL and L2 learners are affected by different types of instruction. Potowski et al. (2009) randomly assigned 127 Spanish HL learners to a processing instruction (PI), traditional instruction (TI), or a tests-only control group. They also assigned 22 L2 learners to either a PI or a TI group, in order to compare the learning outcomes of the two populations. The targeted structure was the Spanish imperfect subjunctive, as used in (2) in adjectival clauses with nonexistent or indefinite referents.

- (2) El año pasado no había políticos que fueran honestos.
 Last year there were no politicians who were-SUBJ honest.

When the referent exists or is definite, as in (3), the imperfect indicative is used.

- (3) El año pasado había políticos que eran honestos.
 Last year there were politicians who were-IND honest.

The PI treatment consisted of explicit rule explanation, a warning against processing strategies that could interfere with learning, and structured input activities. The TI treatment had the same number of total activities and instances of the target form as the PI condition. However, instead of instruction about processing strategies and structured input activities, the TI treatment contained output-focused practice similar to that commonly found in heritage Spanish textbooks, such as *¡Conozcámonos!* (Mrak & Padilla, 2006) and *Nuestro idioma, nuestra herencia* (García, Carney, & Sandoval, 2010).

All learners completed written interpretation, production, and grammaticality judgment tests the day before the instruction and one day after the instruction. Results showed that HL and L2 learners in both PI and TI groups made significant pre–posttest gains on the interpretation and production tests. (There were no gains for the uninstructed control group participants.) On the grammaticality judgment test, only the instructed L2 learners showed significant pre–posttest improvement. Most interesting, perhaps, was the comparison between the instructed HL and L2 groups; L2 learners made larger pretest–posttest gains than their HL learner counterparts. Potowski et al. (2009) concluded that future studies need to investigate how instruction differentially affects HL and L2 learners’ cognitive processing and language development.

Because there was no delayed posttest, it is not possible to make claims about the durability of the effects of instruction. Similarly, the results suggest that HL learners’ acceptability judgments may not be as responsive to instruction as other areas of knowledge, such as controlled production, are. Although the precise reasons for this difference are unclear, it is possible that HL learners rely largely on implicit knowledge, which is more stable and resistant to change, to judge sentences, whereas they draw more on their explicit knowledge for controlled production. Another possible explanation for these findings has to do with HL learners’ reduced exposure to the written form of the HL; all of the acceptability judgment tasks were presented in writing, so they might not provide an accurate assessment of HL learners’ knowledge. Some recent studies have addressed this limitation of written acceptability judgment tests by using a bimodal sentence presentation format with HSSs, whereby stimuli are simultaneously presented both aurally and in writing (Montrul, Bhatt, & Girju, 2015).

In Torres (2013) 34 HL learners and 49 L2 learners were randomly assigned to a control group, or to \pm complex instruction groups on the use of subjunctive or indicative with

Spanish adjectival clauses. All participants completed oral and written production pretests, immediate posttests, and delayed posttests (1–2 weeks after instruction). Each test item consisted of a contextualizing sentence, followed by an incomplete sentence that the participant needed to complete either orally or in writing, depending on the test modality, using an adjectival clause either requiring the present subjunctive or the present indicative.

Control group participants completed the tests but received no instruction, whereas participants in the two complexity groups received computerized task-based instruction on the targeted form that included written feedback. If participants chose the correct verb mood, they saw a message simply saying, “Si/Yes” on the screen. If their mood choice was incorrect, they saw a written recast on the screen (in which the entire sentence with the correct verb form was shown.)

Instructed HL and L2 learners showed comparable gains in oral production, but L2 learners had larger pretest to delayed posttest gains in written production than HL learners. Exit questionnaire responses also suggested that HL and L2 learners may have approached the tasks differently, perhaps based on the context of acquisition. Specifically, L2 learners were more focused on language forms (frequently indicating that they “formed rules” about when to use the subjunctive or indicative during the instruction), whereas HL learners were more focused on content and meaning-making in the tasks. Their comments also indicated that they were less likely to perceive the written recast feedback as corrective, similar to what Gass and Lewis (2007) showed in their study on HL learners’ perception of *oral* corrective feedback.

In the only study to investigate the effects of oral corrective feedback on HL learning outcomes, Kang (2010) used a pretest–posttest/delayed posttest design and divided 45 Korean HL learners into four groups based on whether they received corrective feedback preemptively (before making an error) or reactively (after making an error) and whether they received explicit feedback (including an indication that the utterance was erroneous, an explanation of why, and the correct form) or implicit feedback (in the form of full or partial recasts). She also had a control group, which completed the same communicative activities (a story sequencing and a spot the differences task) targeting the Korean past tense but received back channeling (responses such as “yes” or “uh-huh” to move the interaction forward) instead of corrective feedback. All learners completed a written grammaticality judgment and an oral elicitation task as pretests and posttests. Pretest and immediate posttest comparisons showed no statistically significant differences between explicit and implicit feedback, but they revealed that reactive feedback was significantly more effective than preemptive feedback, and that all experimental groups outperformed the control. Gains were largely maintained at the time of the delayed posttest, 4 weeks after the treatment. Using the same tests and tasks with Korean L2 learners, Kang (2009), found similar results. The high salience of the targeted past-tense form in Kang (2009, 2010) likely played a role in making the recast feedback as effective as the explicit feedback.

Given the small number of studies that have empirically investigated the effects of instruction on HL learners’ linguistic development and the sizable differences among the studies, it is difficult to draw firm conclusions about what aspects of instruction are particularly beneficial. Nevertheless, we can tentatively conclude that HL learners make learning gains from a variety of instructional techniques commonly used in L2 classrooms. By the same token, HL learners’ early exposure to the HL also appears to affect their orientation to the language and to instruction, such that their learning gains often appear not to be as large as those of L2 learners. Future research is needed to gain a more nuanced understanding of

these issues and to determine whether these apparent differences in learning gains are real or whether they are artifacts of the testing instruments.

Is it Beneficial for L2 Learners and HL Learners to Share the Same Classroom?

Having reviewed the limited number of studies investigating the effects of instruction on HL learners' morphosyntactic development, we turn now to studies that have investigated learner–learner interactions, because peer feedback episodes are sites where learning can occur (Bowles & Adams, 2015). Blake and Zyzik (2003) and Bowles (2011) investigated interactions between Spanish L2 and HL learners who were paired in a lab setting. Blake and Zyzik found that text chat-based interaction was more beneficial for L2 learners than their HL counterparts, who had higher Spanish proficiency. Bowles (2011) analyzed the initiation and resolution of language-related episodes (LREs) in the interactions of proficiency-matched HL-L2 pairs who completed both written and oral tasks. She found that both L2 and HL learners initiated a similar number of LREs across oral and written tasks and that the LREs initiated by both types of learners were resolved in equal proportion. Nevertheless, the data revealed different patterns by the two learner types on the written task: 47 of the 70 orthography-focused LREs (67%) were initiated by HL learners, while the other 23 (33%) were initiated by L2 learners, a finding underscoring HL learners' gaps with written language as a result of their language learning experience. Their L2 partners appeared to have a complementary skill set and were able to resolve the HL learners' orthography LREs in a target-like way in more than 90% of cases, underscoring their familiarity with written Spanish, a primary source of input in the classroom.

Bowles, Adams, and Toth (2014) analyzed the task-based interactions of 13 L2-L2 and 13 L2-HL dyads in an intermediate-level, university Spanish-language classroom (13 L2-L2 dyads and 13 L2-HL learner dyads). The study sought to determine whether the dyads differed in their focus on form or in the amount of talk during interaction. Results revealed that the two types of dyads were largely similar, although LREs were more likely to be resolved in a target-like way by L2-HL pairs than by L2-L2 pairs, and there was significantly more target language talk, compared to English use, in mixed pairs. L2 learners used Spanish significantly more with HL learners than they did with other L2 learners, suggesting different conversational norms in the two pair types. Furthermore, posttask perception questionnaire data indicated that L2 and HL learners alike saw the interaction as a greater opportunity for the L2 learners' development than for the HL learners, calling into question whether intermediate language classrooms like this one meet the needs of HL learners. If HLs take classes alongside L2 learners, these data suggest that care should be taken to provide tasks that address the needs and linguistic profiles of both kinds of learners, so that classroom interactions will yield mutual benefits. Indeed, Bowles (2011) suggests that HL and L2 learners can work together for mutual benefit, if oral and written tasks are balanced and proficiency levels are similar. Specifically, L2 learners may benefit from their HL partner's speaking ability in oral tasks, while HL learners may benefit from their L2 partner's greater familiarity with written Spanish in writing tasks.

Finally, Warner (2014) examined the LREs of 9 proficiency-matched Spanish HL-HL dyads who completed the same three collaborative tasks as in Bowles (2011). Transcripts of their interactions revealed 100 LREs focused on orthography (39%), grammar (33%), vocabulary (27%), and pronunciation (1%). The total number of LREs from these HL-HL pairs is significantly lower than the 202 LREs produced by the 9 HL-L2 pairs who completed

the tasks in Bowles (2011). Also, fewer LREs were resolved by the dyads in Warner (2014), and a smaller proportion of the resolutions were target-like. Both the largest percentage of LREs and the greatest number of unresolved LREs focused on orthography, attesting that HL learners struggle with issues such as spelling and accent placement (Montrul, 2008).

All in all, research suggests that HL learners, like L2 learners, benefit from instruction, but it is premature to say with certainty what aspects of instruction are optimal for promoting HL development. When it comes to interaction in the classroom, when learners are at similar proficiency levels and tasks are carefully designed, there can be mutual linguistic benefits for L2 and HL to work together, because their contexts of acquisition confer complementary strengths and weaknesses. By the same token, if there are sizable differences in proficiency between the two learner groups and/or when tasks are not carefully structured to maximize the learning potential for each (as in Bowles et al., 2014), learning opportunities may be unbalanced and often favor L2 learners. That is not to say that such effects of sizable proficiency differences are unique to HL-L2 groups; large proficiency differences could also have a similar impact on L2 learner pairs.

Pedagogical Implications

As our review of the existing research shows, there have been few studies investigating the effects of instruction on HL acquisition and the sizable differences between the studies make drawing firm pedagogical implications difficult. A number of recent articles (Carreira, 2012; Fairclough, 2006), books (Beaudrie, Ducar, & Potowski, 2014), and online modules from the National Heritage Language Resource Center (http://startalk.nhlrc.ucla.edu/Default_startalk.aspx) provide guidance to HL instructors about best practices, drawing not only on this research but also on teachers' experience and research from general education and L1 acquisition.

Teaching Tips

- Know your students' motivations and goals for enrolling in HL classes.
- Use student self-reports and teacher-made or standardized assessments to gauge students' incoming skills in the HL.
- Expect to have multilevel, diverse HL classrooms—which may also enroll L2 learners.
- Use differentiated instruction techniques to meet individual students' needs.
- Focus on all four skills—not just reading and writing. Use a mix of activity types and topics.
- Pair students with complementary strengths and weaknesses together.
- Consider content-based instruction for HL learners with sufficient proficiency.
- Provide corrective feedback using a range of techniques.
- Engage in frequent self-reflection on your attitudes and pedagogical practices.

In any classroom, it is fundamentally important to know your students' goals and motivations. Given the variety of reasons that learners enroll in classes in their HL, this is all the more essential for HL instructors. Common goals for pursuing HL study range from connecting with cultural roots, wanting to be able to communicate with family and friends in the US and abroad, learning to read and write, and even attaining professional-level

proficiency in the four skills for a future career (Carreira & Kagan, 2011). The multifaceted, personal nature of HLs' goals is also illustrated in Montrul (2016, p. 3) whose Spanish and Hindi HSS' collective voices express a desire for expanded vocabulary and grammar and improved speaking and fluency so that they can feel comfortable and confident using their HL.

Instructors can gauge their students' goals and motivations by having them fill out a survey at the beginning of the course and then again at the end, to determine to what extent those goals were met. Teachers can construct their own surveys or use Carreira and Kagan's (2011) survey, or portions thereof, as a model.

Armed with knowledge about students' goals and expectations, HL teachers can tailor their instruction to meet those needs, to the extent practical and possible within the curriculum. Given that HL instruction occurs in a range of classroom contexts, running the gamut from high-school and college language courses to volunteer, community-run language "schools" that meet in public or community spaces like churches or libraries, curricula and the flexibility to modify them will vary.

Furthermore, teachers should get a sense of their learners' incoming level of language ability in the four skills, through a combination of student self-report and instructor assessments (which, depending on the language and the availability of resources, can be formal or informal). Because learners come to the classroom with varying degrees of exposure to their HL, literacy skills, and prior formal instruction, HL instructors should expect their classrooms to be multilevel (Carreira & Kagan, 2011). In this respect, what HL instructors experience in all of their courses is similar to what L2 instructors experience with intermediate and advanced-level learners, who come to the classroom differing widely in knowledge, proficiency, strengths and weaknesses, and prior instruction in the language.

Although some universities have special courses or course sequences for HL learners (at least in some high enrollment languages), at the higher proficiency levels (Beaudrie, 2012), HL and L2 learners are typically enrolled in the same classes. Therefore, in addition to the challenge of multilevel HL classes, instructors may have to contend with classrooms enrolling two different learner profiles. Later we provide concrete suggestions for teachers struggling to meet the needs of such diverse student groups (Henshaw & Bowles, 2015).

Addressing the issue of multilevel HL classes, Carreira (2012) has suggested that HL teachers use differentiated instruction techniques so that the classroom is not a one-size-fits-all learning environment but rather one that is tailored to the needs of individual learners, by having a variety of learning stations throughout the classroom. Henshaw and Bowles (2015) suggest that where available, technology can be effectively leveraged to provide differentiated instruction, with teachers assigning activities to students to give them extra practice in areas they struggle with, without assigning the same activities to everyone.

Although research has repeatedly called attention to HL learners' weaker reading and writing skills, literacy should not be the sole focus of HL instruction. In fact, HL learners have been shown to make oral proficiency gains (moving to "Advanced" and "Superior" levels on the ACTFL speaking scale) as a result of language instruction (Swender, Martin, Rivera-Martinez, & Kagan, 2014). This underscores the need for teachers to incorporate activities that engage the four skills and, at higher levels, to move learners from concrete to more abstract language use, such as that used to support opinions, defend a position, or make hypotheses. In terms of the types of activities used in HL classrooms, teachers should strive for balance, not relying heavily on one activity type (e.g., translation or dictation) or skill but mixing up the types of activities and skills targeted. In pairing students, teachers should aim to match partners with complementary skill sets, to encourage the

potential for mutual benefit to both students (Bowles, 2011; Bowles et al., 2014). As for topics covered in the HL classroom, HL textbooks often focus on themes such as immigration and race, which may be of interest to some learners but also have the potential to alienate US-born HL learners who do not identify personally with the immigrant experience. In addition to addressing such topics, teachers could incorporate other themes of broad interest to students from all backgrounds (e.g., bilingualism, health, technology, social justice, music/art/film).

At advanced levels, HL learners could benefit from extensive HL input in a meaningful, natural setting provided by free voluntary reading (Krashen, 1998) and content-based instruction (CBI). In CBI, the target language is used to teach a particular subject matter and the goal is integrated content and language learning. CBI courses could focus on virtually any subject matter of interest, including topics related to linguistics or dialectal variation (e.g., “Bilingualism,” “Spanish in the US”) or culture (e.g., “Twentieth Century Chicano Films”) and would expose learners to new content and appropriate language with which to discuss it. Thus, CBI courses with substantial reading assignments could increase HL learners’ vocabulary and improve their command of more formal registers, while allowing for the incorporation of both preemptive and reactive focus on form to address problematic structures.

As for corrective feedback in the HL classroom, Ducar (2008) found that an overwhelming majority (91%) of Spanish HL learners wanted to have their errors corrected. This runs counter to the commonly held belief that error correction could discourage learners or decrease their motivation to continue taking HL classes. Although it is premature to say what form of corrective feedback is best for HL learners, it appears that more explicit types of corrective feedback, provided in response to learners’ errors, may be most effective, at least on low salience forms, given that more implicit types of corrective feedback, such as recasts, may go unnoticed or be mistaken as feedback on meaning, rather than form. We suggest that teachers employ a variety of feedback types, including recasts and explicit corrections, and that they be alert to their students’ responses to those feedback moves and willing to adapt their strategies as needed. Beaudrie et al. (2014) discuss the importance of valuing students’ own colloquial language variety as they learn a standard variety of the language. HL teachers must not show disparaging attitudes toward the language that the students bring to the classroom, even when HL learners speak dispreferred or non-prestigious varieties of the HL. Henshaw and Bowles (2015) recommend that HL teachers engage in frequent self-reflection about their own beliefs and resulting behaviors in the language classroom so that they can become aware of any preconceived notions they might have (e.g., “HL learners want an easy A”) and address those before they negatively impact the classroom environment. Henshaw and Bowles (2015) also emphasize the importance of self-reflection for all teachers, regardless of whether they teach L2 learners, HL learners, or a mixture of the two, so that they can assess which strategies are working with a particular class and which need to be modified. In the context of HL teaching, where the research base is small, this is all the more important.

Future Directions

ISLA has existed as its own area of inquiry for several decades, informing both SLA theory and practice, but as an emerging field IHLA is in need of more systematic research. Given the attested differences between HL and L2 learners and the increasing number of

HL learners enrolling in classes in their heritage language, more research on the outcomes of instructed HL acquisition is urgent to guide instructional practices that are maximally effective for HL learners.

In order to advance knowledge in IHLA, questions that need to be addressed in more detail are (1) how classroom instruction is beneficial for HL development and maintenance, (2) what features make such instruction most effective, and (3) whether instruction helps restructure the implicit linguistic system and results in long-lasting gains. As Bowles (in press) has argued, in laying out a research agenda for IHLA, it is not necessary to reinvent the wheel; rather, ISLA research should be taken as a point of departure. First steps toward answering these questions include taking a broad approach, one that includes both descriptive and experimental studies of HL classrooms with learners of different HLs, at a range of proficiency levels and ages. Previous research has focused mainly on university HL learners of Spanish, but future studies should investigate the efficacy of instruction on child and adolescent HL learners of a range of languages, because there may well be important inter-language differences that affect HL maintenance and development. It is critical that future IHLA research not focus just on morphosyntax but also on other language domains, such as vocabulary, semantics, pragmatics, and phonetics/phonology. Additionally, it may be important to consider not only the effects of instruction on specific linguistic targets but also more globally, as well as the effects of instruction on learners' attitudes and on their motivations to continue taking courses in the HL.

Bowles (in press) has argued that IHLA should follow the early trajectory of ISLA, whereby experimental studies should compare instructed groups to matched uninstructed groups to assess the benefits of instruction to determine whether, like in ISLA, instructed HL learners outperform uninstructed (naturalistic) learners. Furthermore, experimental studies should compare beginning and end-of-course outcomes to assess in what areas learners make gains and to gauge the extent of those gains. Studies should also compare the effects of different pedagogical methods on HL learners in order to isolate which features of instruction are most effective for this learner population. In all cases, research should employ both immediate and delayed posttests to establish whether the effects of instruction are durable and should assess both implicit and explicit knowledge. Research is needed in both controlled laboratory environments and in a variety of classroom settings (i.e., classrooms that include HL and L2 learners and classrooms that enroll only HL learners.)

To conclude, as the number of learners who receive classroom instruction in their HL increases, language educators have a responsibility to engage in systematic IHLA research to inform practice. It is our hope that the questions raised in this chapter will spark empirical IHLA research for decades to come.

References

- Albirini, A. (2014). The role of the colloquial varieties in the acquisition of the standard variety: The case of Arabic heritage speakers. *Foreign Language Annals*, 47, 447–463.
- Au, T., Knightly, L., Jun, S.-A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, 13, 238–243.
- Beaudrie, S. (2012). Research on university-based Spanish heritage language programs in the United States: The current state of affairs. In S. Beaudrie & M. Fairclough (Eds.), *Spanish as a heritage language in the United States: State of the field* (pp. 203–221). Washington, DC: Georgetown University Press.

- Beaudrie, S., Ducar, C., & Potowski, K. (2014). *Heritage language teaching: Research and practice*. Columbus, OH: McGraw-Hill Education.
- Blake, R., & Zyzik, E. (2003). Who's helping whom? Learner/heritage-speakers' networked discussions in Spanish. *Applied Linguistics*, 24, 519–544.
- Bowles, M. (2011). Exploring the role of modality: L2—heritage learner interactions in the Spanish language classroom. *Heritage Language Journal*, 8, 30–65.
- Bowles, M. (in press). Outcomes of classroom Spanish heritage language instruction: State of the field and an agenda for the future. In K. Potowski (Ed.), *The Routledge handbook of Spanish as a heritage/minority language*. New York: Routledge.
- Bowles, M., & Adams, R. (2015). An interaction approach to learner–learner interaction in second and foreign language classrooms. In N. Markee (Ed.), *Handbook on classroom discourse* (pp. 198–212). Boston: Wiley.
- Bowles, M., Adams, R., & Toth, P. (2014). A comparison of L2-L2 and L2-heritage learner interactions in Spanish language classrooms. *Modern Language Journal*, 98, 497–517.
- Bowles, M., & Montrul, S. (2009). Instructed L2 acquisition of differential object marking in Spanish. In R. Leow, H. Campos, & D. Lardiere (Eds.), *Little words: Their history, phonology, syntax, semantics, pragmatics and acquisition* (pp. 199–210). Washington, DC: Georgetown University Press.
- Bylund, E. (2009). Maturation constraints and first language attrition. *Language Learning*, 59, 687–715.
- Bylund, E., & Díaz, M. (2012). The effects of heritage language instruction on first language proficiency: A psycholinguistic perspective. *International Journal of Bilingual Education and Bilingualism*, 15, 593–609.
- Carreira, M. (2012). Meeting the needs of heritage language learners: Approaches, strategies, and research. In S. M. Beaudrie & M. Fairclough (Eds.), *Spanish as a heritage language in the United States* (pp. 223–240). Washington, DC: Georgetown University Press.
- Carreira, M., & Kagan, O. (2011). Results of the national heritage language survey: Implications for teaching, curriculum design, and professional development. *Foreign Language Annals*, 44, 40–64.
- Ducar, C. M. (2008). Student voices: The missing link in the Spanish heritage language debate. *Foreign Language Annals*, 41(3), 415–433.
- Fairclough, M. (2006). La enseñanza del español como lengua de herencia: Un curso de preparación para docentes. *Revista Iberoamericana de Lingüística*, 1, 30–50.
- García, H. A., Carney, C., & Sandoval, T. (2010). *Nuestro idioma, nuestra herencia*. New York: McGraw Hill.
- Gass, S., & Lewis, K. (2007). Perceptions of interactional feedback: Differences between heritage language learners and non-heritage language learners. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 79–99). Oxford: Oxford University Press.
- Guijarro Fuentes, P. (2012). The acquisition of interpretable features in L2 Spanish: Personal a. *Bilingualism: Language and Cognition*, 15, 701–720.
- Håkansson, G. (1995). Syntax and morphology in language attrition: A study of five bilingual expatriate Swedes. *International Journal of Applied Linguistics*, 5, 153–171.
- He, A. W. (2006). Toward an identity theory of the development of Chinese as a heritage language. *Heritage Language Journal*, 4, 1–28.
- He, A. W. (2010). The heart of heritage: Sociocultural dimensions of heritage language learning. *Annual Review of Applied Linguistics*, 30, 66–82.
- Henshaw, F. G., & Bowles, M. A. (2015, February). *Meeting the needs of L2 and HL learners in mixed classes without sending mixed messages*. Workshop presented at the II Symposium on Spanish as a Heritage Language, Lubbock, TX.
- Jo, H. (2001). Heritage language learning and ethnic identity: Korean Americans' struggle with language authorities. *Language, Culture and Curriculum*, 14, 26–41.

- Kang, H. (2009). The relative efficacy of explicit vs. implicit feedback in the learning of a less commonly-taught foreign language. *International Review of Applied Linguistics*, 47(4), 303–324.
- Kang, H. (2010). Negative evidence and its explicitness and positioning in the learning of Korean as a heritage language. *Modern Language Journal*, 94(4), 582–599.
- Keating, G. D., VanPatten, B., & Jegerski, J. (2011). Who was walking on the beach? Anaphora resolution in Spanish heritage speakers and adult second language learners. *Studies in Second Language Acquisition*, 33(2), 193–221.
- Krashen, S. (1998). Language shyness and heritage language development. In S. Krashen, L. Tse, & J. McQuillan (Eds.), *Heritage language development* (pp. 15–29). Culver City: Language Education Associates.
- Lightbown, P. M. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 177–196). New York: Cambridge University Press.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic Press.
- Montrul, S. (2008). *Incomplete acquisition in bilingualism: Examining the age factor*. Amsterdam: John Benjamins.
- Montrul, S. (2010). How similar are L2 learners and heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics*, 31, 167–207.
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Montrul, S., Bhatt, R., & Girju, R. (2015). Differential object marking in Spanish, Hindi and Romanian as heritage languages. *Language*, 91(3), 564–610.
- Montrul, S., & Bowles, M. (2010). Is grammar instruction beneficial for heritage language learners? Dative case marking in Spanish. *Heritage Language Journal*, 7, 47–73.
- Mrak, N. A., & Padilla, E. K. (2006). *¡Conozcámonos!: Curso práctico de español para el estudiante bilingüe en los Estados Unidos*. New York: Heinle.
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Oh, J. S., & Au, T. K. (2005). Learning Spanish as a heritage language: The role of sociocultural background variables. *Language, Culture, and Curriculum*, 18, 229–241.
- Peyton, J. K., Carreira, M., Wang, S., & Wiley, T. G. (2008). Heritage language education in the United States. In K. A. King, N. Schilling-Estes, L. Fogle, J. J. Lou, & B. Soukup (Eds.), *Sustaining linguistic diversity* (pp. 173–186). Washington, DC: Georgetown University Press.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.
- Potowski, K. (2002). Experiences of Spanish heritage speakers in university foreign language courses and implications for teacher training. *ADFL Bulletin*, 33, 35–42.
- Potowski, K., Jegerski, J., & Morgan-Short, K. (2009). The effects of instruction on linguistic development in Spanish heritage language speakers. *Language Learning*, 59, 537–579.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Amsterdam: John Benjamins.
- Song, M., O'Grady, W., Cho, S., & Lee, M. (1997). The learning and teaching of Korean in community schools. In Y.-H. Kim (Ed.), *Korean language in America 2* (pp. 111–127). Honolulu, HI: American Association of Teachers of Korean.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, 263–308.
- Swender, E., Martin, C. L., Rivera-Martinez, M., & Kagan, O. E. (2014). Exploring oral proficiency profiles of heritage speakers of Russian and Spanish. *Foreign Language Annals*, 47(3), 423–446.

Silvina Montrul and Melissa Bowles

- Torres, J.R. (2013). *Heritage and second language learners of Spanish: The roles of task complexity and inhibitory control* (Unpublished doctoral dissertation). Georgetown University.
- Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition*, 15, 181–204.
- Warner, S. (2014). *Task-based interactions between Spanish-speaking heritage learners* (Unpublished undergraduate senior thesis). University of Illinois, Urbana-Champaign, IL.
- White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7, 133–161.

Section VI

Instructed Second Language Acquisition Research Methods



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Quantitative Research Methods

Luke Plonsky

Background

Most phenomena addressed by instructed second language acquisition (ISLA) research—explicit instruction, task complexity, linguistic knowledge, for example—are qualitative in nature; that is, most of the constructs we study are not inherently numeric. More often than not, however, ISLA researchers choose to measure and quantify such variables. The same is true in many other areas of applied linguistics, but the domain of ISLA is perhaps unique in that it is so heavily quantitative (Gass, 2009; Plonsky, 2013; cf. De Costa, Valmori, & Choi, this volume). It is in part for this reason—our increasingly strong reliance on quantitative methods (see Loewen & Gass, 2009)—that we must demand high standards of quantitative practices from our colleagues and from ourselves. Although statistics can allow for greater objectivity, systematicity, and ease of analysis, a quantitative approach cannot ensure objectivity, internal validity, or study quality (see Plonsky & Gonulal, 2015).

Another source of urgency for rigor in quantitative ISLA methods stems from the close connection and impact between research in this area and second language (L2) pedagogy. ISLA, as much or perhaps more than any other domain within applied linguistics, is expected to contribute to L2 practice. In order to do so in reliable and meaningful ways, however, the research methods from which our findings are derived must be sound and trustworthy. Without sound methodological practices, the results of ISLA research, the majority of which works with quantitative data, will not be trustworthy and will therefore fail to enhance our understanding of best classroom practices.

Before going further, it might be useful to describe the main features of the “typical” ISLA study. The general design, illustrated in Figure 28.1, is rather straightforward: Researchers examine learning that occurs as the result of one or more types of instruction (i.e., treatments) provided to one or more groups of L2 learners. As will soon become apparent, numerous quantitative and methodological options are available within this basic framework. Several of these options are included in boxes (A)–(D) in Figure 28.1. The arrows pointing to these boxes originate at the element in the design process most relevant to the decisions that must be made. Box (D), for example, includes decisions concerning experimental treatments. Some of the design options presented here are associated with enhanced experimental

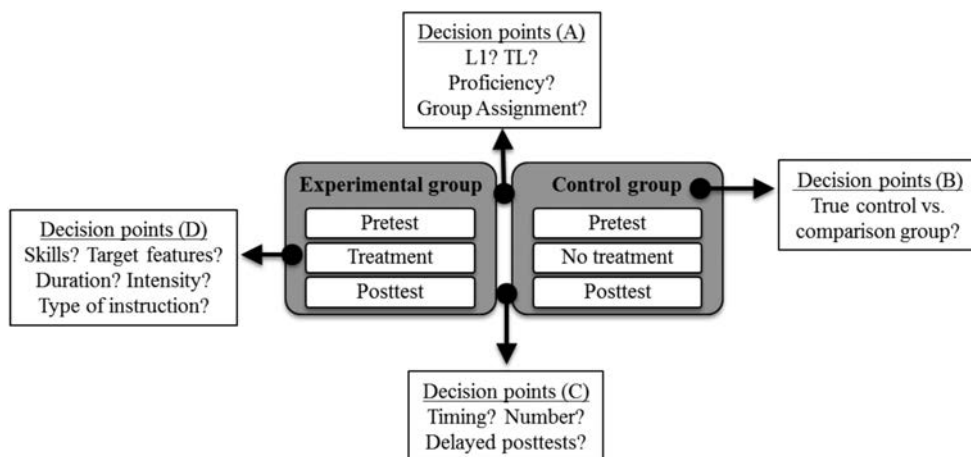


Figure 28.1 Basic design scheme and major decisions in ISLA research

control or quality. However, others are equally viable options that fork into different paths and corresponding results.

One early methodological consideration, highlighted in box (A) of Figure 28.1, involves selecting the participants. Researchers must make decisions, for example, regarding the target language (TL) of interest. Likewise, researchers may determine that only participants with certain first language (L1) backgrounds are eligible to participate, or they may choose not to control for this variable. Still in box (A), proficiency is another background variable that researchers may either control for or at least consider in terms of participant selection. This variable is frequently measured in ISLA research in order to ensure that groups are (roughly) equal prior to receiving a treatment. However, proficiency is notoriously difficult to operationalize and assess (see Hulstijn, 2012; Thomas, 2006; Tremblay, 2011). Standardized tests such as the Test of English as a Foreign Language (TOEFL) may provide high validity and reliability, but they can be expensive and time-consuming to administer. And although such instruments are useful for assessing overall proficiency, ISLA research is often interested in a particular set of target language structures that may or may not be found in a standardized test. In order to overcome these challenges, some scholars employ alternatives such as (1) self-assessments (see Ross, 1998), (2) in-house/researcher-developed tests designed to target general proficiency and/or individual target structures (e.g., Sato & Lyster, 2012), and (3) semesters or years of classroom instruction as a proxy for proficiency (i.e., “seat time”; e.g., Plonsky & Loewen, 2013). Still other studies, such as those interested in incidental focus on form or that do not want to alert learners to the target structures, may choose not to include a pretest in the design (e.g., Loewen, 2005). No single approach or assessment is perfect; rather, each must be evaluated against the goals of the study, the nature of the target structures, and the population to which the study seeks to generalize (Norris & Ortega, 2012).

In true experimental designs, participants are assigned randomly to experimental conditions. In doing so, groups are then often assumed to be of equal proficiency. Random

assignment, which is what distinguishes true experiments from quasi-experiments, is highly valued in interventionist research throughout the social, educational, and medical sciences because it allows the researcher to attribute a causal relationship between group differences and their respective treatment, as opposed to other, preexisting group differences, characteristics, or experiences (see Loewen & Plonsky, 2015; Shadish, Cook, & Campbell, 2002). This practice is ideal and should be applied when feasible within the constraints and goals of the study. That said, in classroom-based ISLA, it is not always possible to assign participants randomly. This potential threat to internal validity is often considered to be offset by the enhanced ecological validity afforded by conducting research in a context that closely resembles those to which the results are meant to be generalized. Finally, it is advisable to collect and compare group pretest data regardless of whether participants are assigned randomly to experimental conditions. One reason for doing so is that with smaller samples it is not appropriate to assume equivalence of groups.

Another assessment-related consideration concerns the number and timing of posttests (see again decisions in box (C) in Figure 28.1), which are included in ISLA research with two purposes in mind. First, posttests are used to compare the performance or outcome (i.e., evidence of learning) of participants in different experimental conditions. Regardless of whether and how many experimental conditions are included in the design, a second function of posttests is to measure what might be called “absolute” gains made over time (as opposed to gains relative to a control group or to another treatment group). Such results provide practical information in the form of an estimate of what learners exposed to a similar treatment might be expected to learn. Providing evidence of “absolute” gains is yet another rationale for including pretests in addition to posttests in ISLA research designs.

A more general point regarding posttesting is that different types of outcome measures often produce different results. For this reason, it is common to find more than one test being used to assess learner performance and/or knowledge. Tanner and Landon (2009), for example, tested the effects of computer-based pronunciation instruction using both a controlled and a spontaneous speech production test. More controlled measures (e.g., cloze passages, multiple choice tests) are usually more straightforward in terms of scoring and therefore less susceptible to rater error; consequently, they also tend to yield larger effects as found in meta-analyses of instruction on L2 pronunciation (Lee, Jang, & Plonsky, 2015) and grammar (Goo, Granena, Yilmaz, & Novella, 2015; Norris & Ortega, 2000). More spontaneous or open-ended outcome measures, by contrast, may possess greater ecological validity as they are more likely to reflect learners’ real-world ability (e.g., Saito, 2012; Saito & Brajot, 2013). However, instruments that involve open-ended production often involve more rater inference and can be more labor-intensive to code different target features.

Researchers are also often interested in measuring the durability of treatment effects by means of one or more delayed posttests. Here as well, it is up to the researcher to determine how many posttests will be administered and how long the delay between them will be. Ellis and He (1999), for example, administered delayed posttests 2, 3, and 4 weeks after their study’s vocabulary intervention. However, it is quite rare to include so many delayed posttests. Overall, 38% (65) of the 172 (quasi)experimental studies in Plonsky’s (2013) sample of quantitative L2 research involved one or more delayed posttests in their design.

As indicated by the items in box (D) in Figure 28.1, choices surrounding the instructional treatment comprise yet another major set of considerations in designing and carrying out (quasi)experimental ISLA research. Among other decisions, the researcher must decide what skill(s) and/or structure(s) will be targeted. Whereas some studies provide instruction on a single form such as the /ɪ/ in L2 English pronunciation (e.g., Saito & Lyster, 2012), others attempt to teach learners a wide variety of structures in a single study. For instance, Macaro and Masterman's (2006) intervention sought to foster L2 French learners' knowledge on features ranging from relative pronouns to past tense formation to prepositions.

ISLA interventions also fall on a number of other continua, such as (1) the complexity of the target features being instructed, (2) the number and length of instructional treatments (akin to the dosage and duration of a medical treatment), (3) the type of instruction (e.g., explicit vs. implicit; comprehension-based vs. production-based), and (4) whether the treatment is provided by the participants' regular classroom teacher or the researcher. Once again, these options are not inherently superior or inferior to each other. Rather, each choice might be reasonable in a given context, providing the study with a unique set of advantages and disadvantages. Some of these choices are also influenced by certain theoretical premises. Furthermore, different types of treatments may also interact with each other, leading to potential impacts on study outcomes. For example, we might expect explicit instruction to be more effective than implicit instruction when target structures are more complex and therefore less likely to be noticed in implicit instruction (see Housen & Simoens, 2016; Spada & Tomita, 2010).

Current Issues

Methodological awareness in ISLA and elsewhere in applied linguistics appears to be increasing, leading to what Byrnes (2013) has referred to as a "methodological turn" (p. 825). For example, in recent years, researchers in the field have (1) introduced novel quantitative techniques such as bootstrapping (Larson-Hall & Herrington, 2010), mixed effect modeling (Cunnings & Finlayson, 2015), and Bayesian data analysis (Mackey & Ross, 2015), (2) questioned and proposed alternatives to null hypothesis significance testing (NHST) and *p* values (e.g., Norris, 2015; Plonsky, 2015a, 2015b), (3) assessed statistical literacy and development (Gonulal, 2016; Loewen et al., 2014), and (4) examined empirically—rather than assuming—the methodological quality in our research (e.g., Liu & Brown, 2015; Plonsky, 2014; Plonsky & Gass, 2011). This section presents a brief discussion of three critical issues currently being discussed as relevant to and employed within the context of quantitative ISLA methods.

Sampling

Recent discussions surrounding this important but often overlooked aspect of ISLA design have address two major issues. The first, statistical power, refers to the likelihood of detecting a statistically significant relationship. When either effect sizes or samples are small, both common in L2 research, power is limited, thereby impeding our ability to derive stable conclusions from our data. Plonsky and Gass (2011), for example, surveyed 174 studies in the interactionist tradition of SLA, coding for a number of study features including sample

sizes and effect sizes (Cohen's d , in this case, a standardized mean difference; see Plonsky, 2012, for a brief introduction). Based on the mean observed sample size ($n = 22$) and the mean effect size ($d = 0.65$), the authors estimated overall post hoc power in the interactionist tradition at .56. This finding indicates that tests of statistical significance in this long-standing area of SLA have been, on average, far below Cohen's (1988) recommendation for minimum power of .80 and only slightly better than chance at appropriately detecting statistical significance. Plonsky's (2013) methodological review of research published in *Language Learning* and *Studies in Second Language Acquisition*, similarly, found post hoc power at .57 (based on a median $n = 19$ and $d = 0.71$). Together these findings provide empirical support, albeit tentative, for warnings voiced from SLA scholars and from related fields (e.g., Cumming, 2012; Oswald & Plonsky, 2010) about the perils of relying on underpowered samples and analyses.

The other major sampling issue currently being addressed is concerned with generalizability of ISLA findings. Whereas low statistical power poses a threat to internal validity and is largely a function of sample size, generalizability (or external validity) is established by means of sampling from (qualitatively) different types of learner populations, contexts, and so forth. To date there is no empirical evidence to support the notion that the field has—or has not—conducted research across many of the contexts and learner demographics that it seeks to generalize to. Anecdotally, however, it appears that it has not. Several scholars have recently criticized the demographic limitations in L2 research, noting especially a lack of research with naturalistic learners, younger children, and adults that vary in socioeconomic status and educational level (e.g., Mackey & Sachs, 2012; Ortega, 2009). As DeKeyser, Alfi-Shabtay, and Ravid (2010) put it, “almost every sample has been one of convenience” (p. 416).

The empirical evidence to support these claims is limited but growing. Liu and Brown's (2015) methodological synthesis of 42 studies of written corrective feedback research revealed, for example, an overreliance on samples of university students (75%) and that 93% of the sample included participants with English as the first or target language. Very similar patterns of participant demographics have also been observed in other subdomains, such as task-based language teaching (Plonsky & Kim, 2016) and learner corpus research (Paquot & Plonsky, in press).

This issue is similar to the observation made about psychological research that often samples English-speaking college students (e.g., Shen et al., 2011) and biomedical research oversampling white males (e.g., Oh et al., 2015), both of which often seek to generalize findings to much broader populations. Although the vast majority of language learning occurs outside of tertiary institutions in North America, much of the L2 research appears to be conducted at US universities. And despite the status of English as likely being the most commonly learned language in the world, there are of course numerous multilingual communities where English is not the target language (e.g., Sridhar, 1994).

The lack of attention to these contexts, learners, and languages introduces a serious threat to the development of both ISLA theory and practice. Theoretically speaking, a comprehensive model for instructed L2 development needs to be able to account for learning that takes place in a variety of contexts and with learners of many different backgrounds. Likewise, for pedagogy, the consequence of limited sampling is that practice cannot be accurately informed and thus these populations of learners cannot be best served (Ortega, 2005).

Statistical Analyses

A good deal of discussion over quantitative analyses in recent years has centered on the relative merits of null hypothesis significance testing (NHST) and p values, on one hand, and practical significance and effect sizes on the other. Quantitative ISLA research, as in much of applied linguistics and throughout the social sciences, has long relied very heavily on p values (mostly resulting from t -tests and ANOVAs) to understand the patterns in our data. This approach provides researchers with an indication of the probability (p) of obtaining the observed differences in mean scores if there is no true difference in the groups' (hypothetical) population means. As in most other social sciences, a 5% probability (p of .05) has traditionally been used as the cutoff for statistical significance. In the case of ISLA, the value is often taken to indicate that one treatment group has outperformed another.

In recent years, however, several applied linguists have echoed colleagues in other fields such as education and psychology who question the utility of p values (e.g., Cohen, 1994; Cumming, 2012), arguing that effect sizes provide more stable and informative estimates of the phenomena and relationships of interest in the field. The three main critiques leveled against p values are presented in contrast to effect sizes such as Cohen's d in Table 28.1 (for recent, full-length discussions, see Norris, 2015; Plonsky, 2015b).

How might these two approaches play out in an actual ISLA study? Perhaps a first question to address concerns the types of analyses found in this domain. ISLA analyses are generally comprised of three main contrasts or tests, indicated by the arrows in Figure 28.2. First, the pretest scores of both groups are compared using a t -test (or ANOVA, if more than two groups are included in the design). If no statistically significant difference is observed, the groups are assumed to be similar and the two remaining analyses can proceed. Second, posttest scores are compared to measure the performance of the control group relative to the experimental group(s) (i.e., a between groups contrast). In Table 28.2, which provides a set

Table 28.1 Three critiques of p -values and corresponding benefits of effect sizes

NHST (p)	Effect sizes (e.g., d , r)
Unreliable; varies in part as a function of sample size	Not dependent on sample size
Uninformative; forces continuous results into a dichotomy	Expresses the magnitude of the relationship in question
Arbitrary convention (.05)	Continuous; can be compared or combined across studies

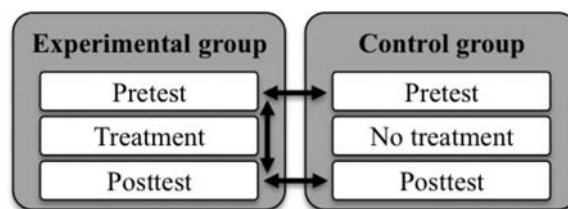


Figure 28.2 Main contrasts in ISLA analyses

Table 28.2 Descriptive statistics for sample study

Group	Pretest		Posttest	
	<i>M (SD)</i>	95% CI	<i>M (SD)</i>	95% CI
Comparison group (<i>n</i> = 20)	<i>a</i> 77 (8)	[73, 81]	<i>c</i> 80 (4)	[78, 82]
Treatment group (<i>n</i> = 20)	<i>b</i> 74 (6)	[70, 78]	<i>d</i> 83 (6)	[80, 86]

of descriptive statistics for this hypothetical scenario, this operation would involve comparing cells *c* and *d*. The experimental and control group's pretest scores can also be compared to their posttests to measure absolute gains (i.e., a within group contrast; cell *b* vs. *d* and *a* vs. *c*, respectively, in Table 28.2).

If we apply this procedure to the data in Table 28.2, we would first run an independent samples *t*-test to compare the groups' pretest scores (again, *a* vs. *b* in Table 28.2). The test reveals that the difference between the two groups is not statistically significant ($t = 1.34$, $p = .19$). A lack of a statistically significant difference can also be found by examining the 95% confidence intervals: each set of intervals includes the other group's mean score. One problem with this (traditional) approach is that a lack of statistical significance is often equated with no difference, a fallacy referred to by Cumming (2012) as the "slippery slope of non-significance" (p. 31). It is useful to remind ourselves that the absence of evidence for a difference is not necessarily evidence of an absence of a difference (see Godfroid & Spino, 2015; Plonsky, 2015a). By contrast, the pretreatment difference between groups can also be expressed as a *d* value of 0.42. Although not a large difference by most standards, it is certainly not a zero or null difference as is often assumed based on a *p* value greater than .05. (For a thorough discussion on interpreting effects sizes in the contexts of L2 research, see Plonsky & Oswald, 2014).

The other main analysis would likely consist of another independent samples *t*-test comparing the group means on the posttest. The *p* value in this case is close to the cutoff for statistical significance but does not cross it ($t = 1.86$, $p = .07$); the NHST-based view of this result would lead the researcher to conclude that there is no difference in the learning that results from the two treatment conditions. Again, however, the effect size for this contrast ($d = 0.59$) tells a different story in which the advantage of the experimental condition might actually be considered sizeable and likely practically or clinically significant. Furthermore, the more nuanced and informative understanding of the pretreatment difference we obtained by examining the effect size also affords us a more appropriate interpretation of the posttest contrast. We can use the effect size for the pretreatment difference between groups as a kind of covariate to adjust for (add to, in this case) the posttreatment difference and thereby obtain a more accurate understanding of the relative gains made in the two treatment conditions: $d = 0.42 + 0.59 = 1.01$. Pretreatment differences may also necessitate an adjustment in the opposite direction. Imagine a study with differences between groups measured as $d = 0.30$ on the pretest and $d = 0.80$ on the posttest, both in favor of the treatment group. It would be inappropriate to simply report the posttest advantage for the treatment group ($d = 0.80$) without accounting for the difference between them that existed prior to the intervention. In this case we would express the adjusted posttest difference by subtracting .30 from .80 to arrive at a *d* value of .50. The only ISLA study I know of that has applied this correction is McManus and Marsden (in press).

I have chosen to illustrate a few common practices and associated potential pitfalls in quantitative ISLA with a sample study involving two groups that were given only two tests: a pretest and a posttest. This design shares many attributes with what is found in most ISLA research but is actually much simpler. It is common for studies in ISLA to include multiple independent, grouping, and/or outcome variables. The reason for this is rather intuitive: language learning and teaching is inherently multivariate in nature (Brown, 2015). Consequently, statistical techniques that are more sophisticated than the *t*-test shown earlier are often needed. It is quite common in ISLA research to examine the effects of multiple treatment conditions with learners at multiple proficiency levels (e.g., beginner, intermediate, advanced) on multiple target features that are tested on multiple instruments. In cases with multiple testing points and/or independent variables, researchers often apply a mixed analysis of variance (ANOVA), which is a type of statistical test that is able to detect not only differences across groups but any interactions between independent variables as well, such as between proficiency level and treatment condition. Because of the developmental nature of much ISLA research, repeated measures ANOVA and mixed repeated measures ANOVA are also used frequently (see Loewen & Plonsky, 2015). When multiple categorical and/or continuous independent variables are included in a design with a single outcome or dependent variable, multiple regression can also be applied. Although it is used much less frequently than ANOVA, the results of multiple regression provide an indication of the amount of variance in the dependent or criterion variable (e.g., learning) that can be accounted for by the predictor or independent variables such as proficiency, L1, treatment length, or feature type (see Plonsky & Oswald, in press).

Data Reporting

Thus far this chapter has discussed a number of issues related to designs and analyses in quantitative ISLA research. Also critical to advancing this domain are the means by which we report and share results. A lack of transparency can negatively affect a domain in multiple ways. Missing or unreported data limits the ability of consumers of primary research to interpret findings. When ISLA researchers fail to report standard deviations, for example, readers are unable to ascertain whether the participants' response to the different treatment conditions was consistent versus highly variable within their respective groups (see Larson-Hall & Plonsky, 2015). At the meta-analytic level, missing standard deviations can also prevent researchers from being able to calculate the study's effect sizes, often forcing the meta-analyst to exclude the study from the sample. Therefore, missing data at the primary level necessarily yields missing data at the meta-analytic level (see additional comments on ISLA meta-analyses later in this chapter).

With these issues in mind and in response to the need for transparency in reports of ISLA research, several meta-analyses and methodological syntheses have coded the primary studies in their samples for whether or not certain types of data were included in published reports (e.g., Mackey & Goo, 2007; Norris & Ortega, 2000; Plonsky, 2011). Also motivating a review of reporting practices are journal and societal guidelines such as the *Publication Manual of the American Psychological Association* (2010), which most L2 journals adhere to and which prescribe, for example, that quantitative studies report full sets of descriptive statistics, effect sizes, confidence intervals, reliability estimates, and so forth.

In the case of standard deviations, reviews in this area paint a rather disappointing picture. Plonsky's (2013) methodological synthesis of 606 quantitative studies published in *Language Learning* and *Studies in Second Language Acquisition* found, for example, that 31% of the sample reported at least one mean without its corresponding standard deviation. Larson-Hall and Plonsky (2015) also examined the extent to which missing standard deviations led to the exclusion of primary studies as reported in 17 meta-analyses of L2 research. The number of studies excluded from meta-analytic samples ranged from just 6% of the final sample (Li, 2010) up to 300% (Wu, 1991). In other words, some meta-analyses actually had to exclude more studies than they were able to include. Considering the generally small samples of primary studies in most meta-analyses of L2 research (see Plonsky & Oswald, 2014), the lack of reporting of standard deviations and other types of data should be considered a serious concern and a potential threat to the validity of meta-analytic findings for the field. If we could assume that data were missing at random, this would be perhaps less of a concern. However, Plonsky's (2013) study also found a tendency to omit descriptive and inferential statistics for analyses that did not achieve statistical significance, a practice that leads to an inflated view of overall effects at the meta-analytic level.

Another practice examined in recent syntheses of L2 research is the reporting of reliability estimates. As with standard deviations, a lack of such data constrains the field in multiple ways. At the primary level, when reliability estimates are not available, readers are unable to gauge the amount of error present in the data. Future research using similar instruments is also left without a guide for what might be expected with a different sample. And at the secondary level, when reliability estimates are not reported, meta-analysts are not able to correct for the attenuation (reduction) of effect sizes that results from measurement error (Plonsky & Oswald, 2015; see also Hunter & Schmidt, 2014). Similar to Larson-Hall and Plonsky's (2015) assessment of the reporting of standard deviations, Plonsky and Derrick (2016) examined the extent to which reliability estimates were found across a number of domains of L2 research that had been subject to meta-analysis or methodological synthesis. Their results were mixed: The presence of reliability estimates was found to range from 6% in studies of L2 practice effects (Nekrasova & Becker, 2009) to 64% in L2 interaction (Plonsky & Gass, 2011; see also Derrick, 2016, and Plonsky & Derrick, 2016, for a meta-analysis of and guide to interpreting reliability coefficients in L2 research).

Data visualization techniques can also be useful for presenting the findings of quantitative ISLA research. A majority of studies in the field currently make use of graphic, nontabular displays, a pattern that appears to be increasing over time (Plonsky, 2014). Unfortunately, we may not be doing so using the most effective or efficient techniques. Hudson (2015) reviewed data visualization practices in 136 empirical studies published in five major L2 journals: *Applied Linguistics*, *Language Learning*, *Modern Language Journal*, *Studies in Second Language Acquisition*, and *TESOL Quarterly*. His findings echo concerns expressed by other scholars such as Larson-Hall and Herrington (2010) and Larson-Hall and Plonsky (2015). In particular, we rely primarily on line graphs and bar graphs, neither of which generally provide a data-rich perspective. Although these types of graphs are easy to create and to interpret, they usually provide little or no information about the dispersion of scores around the mean. Hudson (2015) and Larson-Hall and Plonsky (2015) recommend the use of more data-rich techniques, whenever possible, such as box plots and scatter plots.

Key Concepts

Causation: A change (in learning, for example) that can be directly attributed to an instructional intervention.

Comparison group: A group of participants that receives a minimal or traditional treatment and whose test data are compared to those of the treatment group(s).

Confidence interval: A range of values within which the true population value is expected to fall, within a given level of probability (usually 95%).

Control group: A group of participants that receives no treatment but that provides test data to be compared with the treatment group(s).

Delayed posttest: A measure of learning administered at some interval following an immediate posttest.

Ecological validity: The fit between the context and procedures of a study and the applied setting in which the results of the study might be generalized, such as an L2 classroom.

Effect size: A quantitative measure of the impact of an intervention or of the magnitude of a relationship, most often expressed as d , r , r^2 , or R^2 .

Experimental design: An interventionist study in which participants are assigned randomly (rather than out of convenience) to experimental conditions.

External validity: The extent to which the findings of a given study can be generalized to other contexts, samples, target features, and so forth.

Intervention: The treatment in a (quasi)experimental study.

Null hypothesis significance testing (NHST): A practice in which statistical results are deemed to be significant (or not) on the basis of the probability (p) of the observed data given no relationship or difference between groups in the true (hypothetical) population.

Posttest: A posttreatment assessment designed to measure learning that takes place as a result of an intervention.

Pretest: A pretreatment assessment administered to ensure (or account for a lack of) comparability of groups and to compare to a posttest to measure gains.

Quasi-experimental design: An interventionist study in which participants are not assigned randomly to experimental conditions.

Statistical power: Within the NHST framework, a statistic that indicates the likelihood of obtaining a statistically significant result if such a relationship or effect exists in the population.

Treatment group: A group of participants that receives an experimental treatment (e.g., instruction).

Empirical Evidence

This handbook is a testament to the vast wealth of knowledge that has accumulated in ISLA (see also Loewen, 2015). In order to make sense of and synthesize the large bodies of quantitative results in this domain, ISLA researchers have turned increasingly to meta-analysis in recent years (Norris & Ortega, 2010; Plonsky & Oswald, 2015).

Figure 28.3 and Table 28.3 present a summary of meta-analytic findings from ISLA across four subdomains: grammar, vocabulary, pragmatics, and pronunciation. Two major findings stand out immediately. First of all, considering the median d value of 1.01 across the set of meta-analyses, instructional interventions in ISLA research generally lead to substantial evidence of learning (see Plonsky & Oswald, 2014, for a guide to interpreting effect sizes

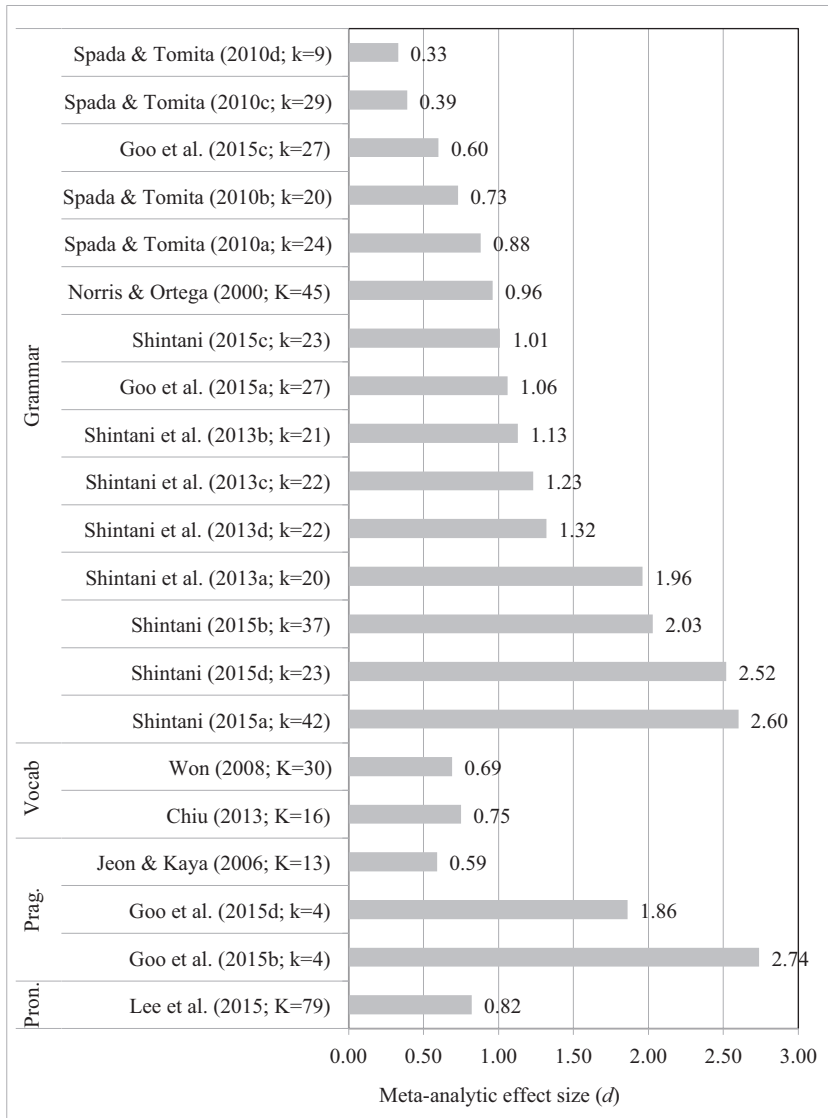


Figure 28.3 Summary of meta-analytic effects in ISLA across subdomains

Notes: K = Total number of studies included in the meta-analysis; k = a subset of the total number of studies included in the meta-analysis. Several meta-analyses presented subgrouped (set) results based on one or more study/treatment features, rather than a single, overall effect.

in the context of L2 research). The second main finding here is that meta-analytic effects vary widely both across and within the different target areas. Within grammar alone, meta-analytic effects range from what would generally be considered small ($d = 0.33$; implicit instruction on simple structures, Spada & Tomita, 2010) to very large ($d = 2.60$; processing instruction, receptive outcome measure, Shintani, 2015). A similarly wide range of meta-analytic effects can be seen for instruction on L2 pragmatics. By contrast, meta-analytic effects for vocabulary instruction, although fewer in number, appear more consistent.

Table 28.3 Descriptors for meta-analyses with multiple groups presented in Figure 28.3

Study	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
Spada and Tomita (2010)	explicit instruction, complex target feature	explicit, simple	implicit, complex	implicit, simple
Goo et al. (2015)	explicit instruction, grammar	explicit, pragmatics	implicit, grammar	implicit, pragmatics
Shintani, Li, and Ellis (2013)	comprehension-based instruction on receptive dependent measure	production-based instruction, receptive	comprehension-based instruction, productive	production-based instruction, productive
Shintani (2015)	processing instruction, receptive outcome measure	processing instruction, productive outcome measure	production-based instruction, receptive outcome measure	production-based instruction, productive outcome measure

In-depth consideration of what might lead to such a wide range of results within individual subdomains is outside the scope of this chapter. However, the causes are likely due in part to the unique operational definitions of each meta-analysis. Plonsky and Brown (2015) examined the results of 18 syntheses of corrective feedback, another domain with results that vary widely at the meta-analytic level, from negative and small ($d = -0.16$ in Truscott, 2007) to positive and large ($d = 1.16$ in Russell & Spada, 2006). Their discussion pointed to a number of different substantive and methodological choices that led to the disparity in results including (1) the meta-analysts' definition of the domain in question and (2) the search techniques employed to obtain studies that fell within that domain.

Future Directions

Quantitative ISLA has made great strides since the field's inception (Loewen & Gass, 2009). However, this chapter has described several methodological concerns that remain and that pose threats to our understanding of instructed L2 development. With these concerns in mind, I close the chapter with two brief sets of recommendations for future research. The first set concerns methodologically oriented studies, many of which were referenced earlier. Work in this area is gaining traction, but it has yet to yield the influence on research practices needed in order to prompt major change in the field. The second set of recommendations is directed toward the work of primary researchers and is based on many of the principles of sound research practice that are described in this chapter. These comments concern designs, analyses, and reporting practices.

In a spirit similar to the suggestions laid out in this section, two recent sets of journal guidelines have been made available that will likely be of great interest both for and beyond the journals they were developed for. Norris, Plonsky, Ross, and Schoonen (2015) provide a detailed prescription for conducting and reporting on quantitative L2 research. Although the guidelines were commissioned by and published in *Language Learning*, a broader reach could certainly be applied. *TESOL Quarterly* published a work with a similar intent quite recently as well, authored by Mahboob et al. (2016). The methodological scope of this set

of guidelines, however, was decidedly broader than those written for *Language Learning*, a choice that might be interpreted as a reflection of the editorial preferences and culture of *TESOL Quarterly*.

Methodologically Oriented Research

In the realm of methodologically oriented research, two types of studies are particularly needed. The first builds on the growing body of methodological syntheses. Studies in this vein can focus on specific research practices (e.g., the use of a particular design feature or statistical technique) within a broad cross-section of L2 research, as in Plonsky and Gonulal's (2015) evaluation of the use of factor analysis, or they can assess methodological practices in a given substantive domain (e.g., written corrective feedback in Liu & Brown, 2015). In either case, systematicity and adherence to synthetic best practices is critical. These types of studies are key to providing precise and empirically supported guidance to future research in ISLA or any discipline (see Ioannidis, Fanelli, Dunne, & Goodman, 2015).

The second major type of methodologically oriented research currently needed involves gaining a better appreciation of the field's statistical literacy and development. We are beginning to understand methodological practices as observed in published studies, but we have very little data to indicate what researchers know, what they are able to do, where and when this knowledge originates, and how it might change over time. Some scholars have called for additional methodological training in graduate programs, but it is not clear what such curricula would or should consist of (Gonulal, 2016; Loewen et al., 2014; Plonsky, 2014).

Improving Future Quantitative ISLA Research

Designs in quantitative ISLA will be stronger and more likely to contribute in meaningful ways to L2 theory and practice with: (1) larger samples; (2) samples consisting of under-researched demographics (e.g., children, older adults, low-literacy learners, learners without English as an L1 or L2); (3) more pretesting; and (4) more delayed posttesting.

With respect to quantitative analyses in ISLA, two complementary approaches are needed. One involves focusing less on the flawed practice of null hypothesis significance testing and more on a thorough consideration and interpretation of descriptive statistics including effect sizes and confidence intervals. The results from these studies can then be brought together and aggregated via research synthesis and meta-analysis. The other approach involves recognizing the multivariate nature of language learning and teaching and using analyses that can simultaneously account for the many concurrent relationships inherent in our data. Put more simply, more multivariate analyses are needed. Primary ISLA studies routinely report the results of dozens of univariate and bivariate analyses in a single study. Oftentimes, a single procedure could address the same questions or set of questions, providing a result that (1) is more parsimonious, (2) preserves experiment-wise power, and (3) will be more in line with the multivariate relationships being addressed (Brown, 2015).

I would also encourage ISLA researchers to consider the potential applicability of novel analytical techniques to the questions they address. These include, for example, mixed effects modeling, Bayesian data analysis, and bootstrapping.

Finally, the recommendations for improving data reporting practices are quite straightforward and can be distilled into four main points. First, full sets of descriptive statistics should be reported for all variables of interest. Second, the results for all analyses worth conducting should be included in the published form or in an online appendix; failing

to do so leads to an inflated view of overall effects at the meta-analytic level. Third, data-rich visualization techniques should be adopted whenever possible. And fourth and last, as required by the journals of the American Psychological Association, I would encourage researchers to share their instruments (e.g., via IRIS) and data with others interested in replication, reanalysis, and meta-analysis, when requested to do so. Better yet, researchers can make such materials available preemptively by uploading them to the IRIS database for research into second language learning and teaching (<http://iris-database.org>; see Marsden, Mackey, & Plonsky, 2016), which, as of this writing, contains over 1,000 tools from published L2 research. Thinking and acting collaboratively and synthetically, rather than individualistically, is key to advancing our understanding of instructed L2 development.

References

- American Psychological Association. (2010). *Publication manual of the American psychological association* (6th ed.). Washington, DC: Author.
- Brown, J. D. (2015). Why bother learning advanced quantitative methods in L2 research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 9–20). New York: Routledge.
- Byrnes, H. (2013). Notes from the editor. *Modern Language Journal*, 97, 825–827.
- Chiu, Y.-H. (2013). Computer-assisted second language vocabulary instruction: A meta-analysis. *British Journal of Educational Technology*, 44, 52–56.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 97–1003.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.
- Cummings, I., & Finlayson, I. (2015). Mixed effects modeling and longitudinal data analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 159–181). New York: Routledge.
- DeKeyser, R., Alfi-Shabtay, I., & Ravid, D. (2010). Cross-linguistic evidence for the nature of age effects in second language acquisition. *Applied Psycholinguistics*, 31, 413–438.
- Derrick, D. J. (2016). Instrument reporting practices in second language research. *TESOL Quarterly*, 50, 132–153.
- Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21, 285–301.
- Gass, S. (2009). A survey of SLA research. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 3–28). Bingley: Emerald.
- Godfroid, A., & Spino, L. (2015). Reconceptualizing reactivity of think-alouds and eye tracking: Absence of evidence is not evidence of absence. *Language Learning*, 65, 896–928.
- Gonulal, T. (2016). *Statistical literacy among second language acquisition graduate students* (Unpublished doctoral dissertation). Michigan State University, East Lansing.
- Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). Amsterdam: John Benjamins.
- Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition*, 38, 163–175.
- Hudson, T. (2015). Presenting quantitative data visually. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 78–105). New York: Routledge.
- Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15, 422–433.

- Hunter, F. L., & Schmidt, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Thousand Oaks, CA: Sage.
- Ioannidis, J.P.A., Fanelli, D., Dunne, D. D., & Goodman, S. N. (2015). Meta-research: Evaluation and improvement of research methods and practices. *PLOS Biology*, 13, e1002264.
- Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 165–211). Philadelphia: John Benjamins.
- Larson-Hall, J., & Herrington, R. (2010). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, 31, 368–390.
- Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65, 127–159.
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36, 345–366.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing*, 30, 66–81.
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, 361–386.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Loewen, S., & Gass, S. (2009). The use of statistics in L2 acquisition research. *Language Teaching*, 42, 181–196.
- Loewen, S., Lavolette, B., Spino, L., Papi, M., Schmidtke, J., Sterling, S., & Wolff, D. (2014). Statistical literacy among applied linguists and second language acquisition researchers. *TESOL Quarterly*, 48, 360–388.
- Loewen, S., & Plonsky, L. (2015). *An A-Z of applied linguistics research methods*. New York: Palgrave.
- Macaro, E., & Masterman, L. (2006). Does intensive explicit grammar instruction make all the difference? *Language Teaching Research*, 10, 297–327.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–451). New York: Oxford University Press.
- Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning*, 62, 704–740.
- Mackey, B., & Ross, S. J. (2015). Bayesian informative hypothesis testing. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 329–345). New York: Routledge.
- Mahboob, A., Paltridge, B., Phakiti, A., Wagner, E., Starfield, S., Burns, A., Jones, R. H., & De Costa, P. I. (2016). TESOL Quarterly research guidelines. *TESOL Quarterly*, 50, 42–65.
- Marsden, E., Mackey, A., & Plonsky, L. (2016). Breadth and depth: The IRIS repository. In A. Mackey & E. Marsden (Eds.), *Instruments for research into second languages: Empirical studies advancing methodology* (pp. 1–21). New York: Routledge.
- McManus, K., & Marsden, E. (in press) L1 explicit instruction can improve L2 online and offline performance: An exploratory study. *Studies in Second Language Acquisition*.
- Nekrasova, T., & Becker, T. (2009). *Effectiveness of practice: A research synthesis and quantitative meta-analysis*. Unpublished manuscript.
- Norris, J. M. (2015). Statistical significance testing in second language research: Basic problems and suggestions for reform. *Language Learning*, 65, 97–126.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Norris, J. M., & Ortega, L. (2010). Timeline: Research synthesis. *Language Teaching*, 43, 61–79.
- Norris, J. M., & Ortega, L. (2012). Assessing learner knowledge. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 573–589). London: Routledge.

- Norris, J.M., Plonsky, L., Ross, S.J., & Schoonen, R. (2015). Guidelines for reporting quantitative methods and results in primary research. *Language Learning*, 65, 470–476.
- Oh, S.S., Galanter, J., Thakur, N., Pino-Yanes, M., Barcelo, N.E., White, M.J., . . . Burchard, E.G. (2015). Diversity in clinical and biomedical research: A promise yet to be fulfilled. *PLOS Medicine*, 12, e1001918.
- Ortega, L. (2005). For what and for whom is our research? The ethical as transformative lens in instructed SLA. *Modern Language Journal*, 89, 427–443.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- Oswald, F.L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85–110.
- Paquot, M., & Plonsky, L. (in press). Quantitative research methods and study quality in learner corpus research. *International Journal of Learner Corpus Research*.
- Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning*, 61, 993–1038.
- Plonsky, L. (2012). Effect size. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 200–202). New York: Routledge.
- Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, 35, 655–687.
- Plonsky, L. (2014). Study quality in quantitative L2 research (1990–2010): A methodological synthesis and call for reform. *Modern Language Journal*, 98, 450–470.
- Plonsky, L. (2015a). Quantitative considerations for improving replicability in CALL and applied linguistics. *CALICO Journal*, 32, 232–244.
- Plonsky, L. (2015b). Statistical power, *p* values, descriptive statistics, and effect sizes: A “back-to-basics” approach to advancing quantitative methods in L2 research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 23–45). New York: Routledge.
- Plonsky, L., & Brown, D. (2015). Domain definition and search techniques in meta-analyses of L2 research (Or why 18 meta-analyses of feedback have different results). *Second Language Research*, 31, 267–278.
- Plonsky, L., & Derrick, D.J. (2016). A meta-analysis of reliability coefficients in second language research. *Modern Language Journal*, 100, 538–553.
- Plonsky, L., & Gass, S. (2011). Quantitative research methods, study quality, and outcomes: The case of interaction research. *Language Learning*, 61, 325–366.
- Plonsky, L., & Gonulal, T. (2015). Methodological synthesis in quantitative L2 research: A review of reviews and a case study of exploratory factor analysis. *Language Learning*, 65, 9–36.
- Plonsky, L., & Kim, Y. (2016). Task-based learner production: A substantive and methodological review. *Annual Review of Applied Linguistics*, 36, 73–97.
- Plonsky, L., & Loewen, S. (2013). Focus on form and vocabulary acquisition in the Spanish L2 classroom. *Language, Interaction, and Acquisition*, 4, 1–24.
- Plonsky, L., & Oswald, F.L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912.
- Plonsky, L., & Oswald, F.L. (2015). Meta-analyzing second language research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 106–128). New York: Routledge.
- Plonsky, L., & Oswald, F.L. (in press). Multiple regression as a flexible alternative to ANOVA in L2 research. *Studies in Second Language Acquisition*.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experimental factors. *Language Testing*, 15, 1–20.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J.M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Philadelphia: John Benjamins.
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 46, 842–854.

- Saito, K., & Brajot, F. (2013). Scrutinizing the role of length of residence and age of acquisition in the interlanguage pronunciation development of English /r/ by late Japanese bilinguals. *Bilingualism: Language and Cognition*, 16, 847–863.
- Saito, K., & Lyster, R. (2012). Effects of form focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning*, 62, 595–633.
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34, 591–626.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Shen, W., Kiger, T. B., Davies, S. E., Rasch, R. L., Simon, K. M., & Ones, D. S. (2011). Samples in applied psychology: Over a decade of research in review. *Journal of Applied Psychology*, 96, 1055–1065.
- Shintani, N. (2015). The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: A meta-analysis. *Applied Linguistics*, 36, 306–325.
- Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning*, 63, 296–329.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60, 263–308.
- Sridhar, S. N. (1994). A reality check for SLA theories. *TESOL Quarterly*, 28, 800–805.
- Tanner, M. W., & Landon, M. M. (2009). The effects of computer-assisted pronunciation readings on ESL learners' use of pausing, stress, intonation, and overall comprehensibility. *Language Learning & Technology*, 13, 51–65.
- Thomas, M. (2006). Research synthesis and historiography: The case of assessment of second language proficiency. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 279–298). Philadelphia: John Benjamins.
- Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research: “Clozing” the gap. *Studies in Second Language Acquisition*, 33, 339–372.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16, 255–272.
- Won, M. (2008). *The effects of vocabulary instruction on English language learners: A meta-analysis* (Unpublished doctoral dissertation). Texas Tech University, Lubbock, TX.
- Wu, C. (1991). *A meta-analysis of intensive audio-lingual language programs: A meta-analysis of intensive audio-lingual language programs: Implications for English as a foreign language instruction in China* (Unpublished doctoral dissertation). University of San Francisco, San Francisco.

Qualitative Research Methods

Peter I. De Costa, Lorena Valmori, and Ina Choi

Background

Qualitative research today, according to Creswell (2012), involves closer attention to “the interpretive nature of inquiry and situating the study within the political, social, and cultural context of the researchers, and the reflexivity or ‘presence’ of the researchers in the accounts they present” (p. 45). Put simply, qualitative researchers themselves constitute a central component of their work, which itself is deeply embedded within a broader social milieu. Also acknowledging the importance of social context but writing specifically in relation to SLA, Friedman (2012) adds that qualitative research is cyclical and characterized by the following features: open inquiry, inductiveness, descriptiveness and interpretiveness, multiple perspectives, and focus on the particular. It is these concerns that have guided much of the explosion in qualitative research in SLA the last 20 years. In their survey of qualitative research published in language teaching and learning journals between 1997 to 2006, Benson, Chik, Gao, Haung, and Wang (2009) observed a 22% increase in qualitative research, compared to just a 10% increase observed by Lazaraton (2000) a decade earlier.

Fueling this growing interest in qualitative SLA research perhaps is the increasing theoretical and methodological diversity within our field. Theoretically, since Firth and Wagner’s (1997) seminal call to recognize the social dimensions of language learning, SLA has witnessed a social turn (Block, 2003), which has resulted in a reconceptualization of the field that now covers a range of alternative approaches (Atkinson, 2011) that include complexity theory, Vygotskian sociocultural theory, language socialization, identity, and conversation analysis. More importantly, the field has also seen encouraging attempts to bridge the social and cognitive divide, as observed in Hulstijn et al. (2014), and most recently in the work of the Douglas Fir Group (2016). In particular, the framework put forward by the group (Figure 29.1) is likely to spur further growth in qualitative SLA because it takes an ecological perspective on how language learning within a classroom (meso level) is shaped by cognitive processes within the individual learner (micro level) and the wider social context (macro level).

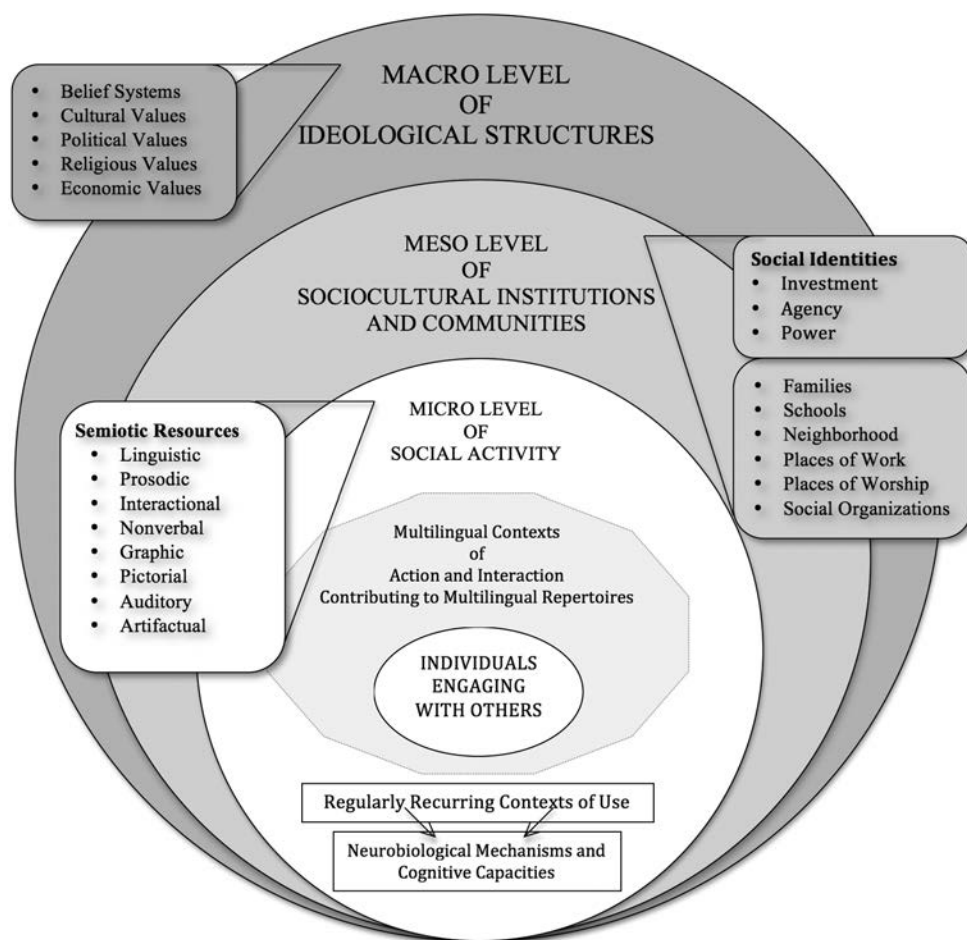


Figure 29.1 Multifaceted nature of language learning and teaching

Source: The Douglas Fir Group, 2016

Methodology books (e.g., Dörnyei, 2007; Hinkel, 2005, 2011; Mackey & Gass, 2015; Richards, 2003) and special issues of journals have also helped raise the visibility and use of qualitative research in SLA. With regard to the latter, the field has benefited from in-depth discussions of interviews (Talmy & Richards, 2011), classroom discourse microanalyses (Zuengler & Mori, 2002), gesture and SLA (Gullberg & McCafferty, 2008), classroom talk (Markee & Kasper, 2004), as well as comprehensive reviews of specific methodologies such as narrative inquiry (Benson, 2014), conversation analysis (Kasper & Wagner, 2014), critical discourse analysis (Lin, 2014), and case study (Duff, 2014). Crucially, qualitative SLA researchers have further benefited from the guidelines (Chapelle & Duff, 2003; Mahboob et al., 2016) published in *TESOL Quarterly*, which also runs a regular and separate Research Issues section. While the qualitative segment of the 2003 *TESOL Quarterly* guidelines focused on case study, conversation analysis, and critical ethnography, the 2016 guidelines

include a discussion of research ethics and provide direction on how to conduct ethnographic research, discourse analysis, and practitioner research. Equally encouraging has been support from SLA-related conferences such as the Second Language Research Forum (SLRF) and the American Association for Applied Linguistics (AAAL) that provide preconference workshops on qualitative research. For the first time in 2016, AAAL also featured a new strand, *Research Methods*, that allows SLA researchers to share their work with their peers with a particular focus on the methods used and also to tap the expanding interest in mixed methods research.

(Re)interpreting ISLA

Collectively, the aforementioned positive developments illustrate the vibrancy of qualitative research in SLA. However, before we proceed any further, we would like to articulate how we interpret ISLA for the purpose of this chapter. More often than not, the classroom context is a key identifying marker in distinguishing ISLA from naturalistic SLA (Loewen, 2015; Pica, 2012; Spada & Lightbown, 2012). As observed by Loewen (2015), “the prototypical context of ISLA is the L2 classroom” (p. 5). Further, he maintains that the main aim of ISLA is to “understand how the systematic manipulation of the mechanisms of learning and/or the conditions under which they occur enable or facilitate the development and acquisition of a language other than one’s own” (p. 3). While we acknowledge of the importance of situating ISLA within a classroom context in this chapter, like the Douglas Fir Group (2016), we also assert that learners within the classroom and the instruction they receive are shaped by broad social processes not only within the classroom but outside as well. By adopting such a multifaceted perspective, we argue that qualitative SLA researchers are in a unique position to better understand not just the mechanisms and conditions of learning but also the dynamics surrounding L2 learners (see Figure 29.1). Such social dynamics also include the language teacher, who plays a significant role in shaping the language development of her students.

In addition, and in line with Loewen (2015), we problematize the notion of the classroom by including virtual classrooms, self-study, and study abroad, because increasingly, much language learning takes place outside of a brick and mortar classroom. Finally, instead of focusing only on language, we prefer to use the notion of *semiotics* (linguistic and nonlinguistic resources such as prosodic, interactional, and nonverbal) because every learner has a rich basket of semiotic resources (see Douglas Fir Group, 2016) available to them. In sum, in taking such a stance on ISLA and through the examples that are discussed in this chapter, we demonstrate how contemporary qualitative ISLA research associated with classrooms provides a helpful lens in understanding how language learning outcomes can be enhanced.

Current Issues

Before we discuss five issues concerning qualitative research, we would first like to make a distinction between methodology and methods. In essence, methodology is the theoretical and paradigmatic lens through which qualitative researchers choose to better understand reality, while methods are the actual instruments and procedures they use to collect their data.

Key Concepts

Reflexivity: The researchers' process of reflecting on their role in the social, cultural, and relational research context by problematizing their assumptions, preconceptions, selection of participants and research setting, and framing of questions.

Semiotics: Resources used by learners, which include linguistic, prosodic, interactional, nonverbal, graphic, pictorial, auditory, and artifactual resources.

Multifaceted nature of language learning and teaching (Douglas Fir Group, 2016): Conceptualization of L2 learning and teaching as an ecological process that takes place across three dimensions: the individual (micro level), the sociocultural context (meso level), and ideological structures (macro).

Methodology and Methods

Examples of qualitative methodology include case study (Duff, 2014), ethnography (Starfield, 2015), conversation analysis (Kasper & Wagner, 2014), and narrative inquiry (Benson, 2014). For succinct but helpful descriptions of various quantitative and qualitative methodologies, see Paltridge and Phakiti (2015). It is important to note that methodologies are informed by the researcher's epistemology and ontology (which are discussed next) and the theoretical framework that guides the study. Further, methodologies are also embedded in rich disciplinary traditions and histories; the intellectual lineage of ethnography, for example, can be traced back to anthropology (Starfield, 2015). More recently, however, ethnography has been used to study classroom cultures. Also important to note is that it is not an uncommon practice for SLA researchers to combine methodologies; for example, De Costa (2015) combined ethnography and case study in his work, while Talmy (2011) combined ethnography with conversation analysis.

Qualitative methodologies are generally constituted by the use of multiple qualitative methods such as interviews and classroom observations (for a discussion of individual methods, see Friedman, 2012; Mackey & Gass, 2015; Richards, 2003). Briefly, on a logistical level, interviews may be structured (i.e., scripted) or semi-structured (i.e., partially scripted), while observations may be closed (i.e., predefined categories for the observation schedule) or open (categories emerge during the observation). These methods are often used in conjunction with data sources such as classroom interactions (student–student or student–teacher, which could be audio- and/or videorecorded), and artifacts (resources that evidence and support learning, such as diaries, and teaching materials). In order to investigate language learner anxiety in a secondary English classroom in Singapore, De Costa (2015), for example, used a combination of these methods and data sources in his year-long ethnographic case study. To get a holistic understanding of his focal students' anxiety, he conducted and audiorecorded semi-structured interviews with them, their peers, and their teachers. Field notes were taken during his open classroom observations, from which analytic categories later emerged. He also used a combination of audiorecording (pair and group work) and videorecording (classwide discussions) in order to examine how classroom interactions may have affected his focal students' anxiety. Artifacts such as graded samples of written work were also collected and subsequently analyzed.

Unpacking Paradigms: The Importance of Epistemology and Ontology

As noted, the research paradigms of ISLA researchers often determine the methodologies they adopt. Distinguishing between ontology (a researcher's views on the nature of reality) and epistemology (a researcher's views on the nature of knowledge and how it can be acquired), Denzin and Lincoln (2011) point out that one's methodology is the research approach used to investigate reality (for a further discussion of research paradigms in applied linguistics, see Phakiti & Paltridge, 2015). Applying this distinction to ISLA research (see Table 29.1), one could argue that two primary paradigms—postpositivism and postmodernism—appear to have shaped the ISLA research agenda.

Building on this paradigmatic distinction, Friedman (2012) rightly points out that “quantitative and qualitative approaches do not map neatly onto postpositivist (quantitative) and postmodern (qualitative) paradigms; qualitative research in SLA,” she adds, “has been and continues to be conducted from a postpositivist perspective” (p. 181). The truth of this observation is borne out in Gurzynski-Weiss and Baralt's (2014) study, which explored learner perception and use of task-based interactional feedback in computer-mediated and face-to-face modes. Using stimulated recall, a form of interview, they sought to (among other things) test the external validity of their experiments by investigating if learners (1) perceived feedback provided during task-based interaction, and (2) recognized the target of the feedback provided during task-based interaction. Thus, interviews—a qualitative method—in the form of stimulated recalls were used to test hypotheses, thereby highlighting the postpositivist paradigm to which the researchers subscribed and illustrating how they adopted a form of *dominance design* (where in this case more emphasis was given to quantitative methods over qualitative methods), as described by Phakiti and Paltridge (2015).

Figure 29.2 is a visual representation of *how* the approach used to investigate reality (i.e., methodology) should be aligned with the paradigm (i.e., researcher's epistemology and

Table 29.1 Paradigms, epistemology, and ontology

ISLA Paradigms	Epistemology (what constitutes knowledge)	Ontology (view of reality)
Postpositivist	Test hypotheses or look for cause–effect relationships in language learning	Seek an objective reality and a single truth
Postmodernist	Try to understand the experiences, abilities, perceptions, and performances of language learners	Seek subjective realities and multiple truths

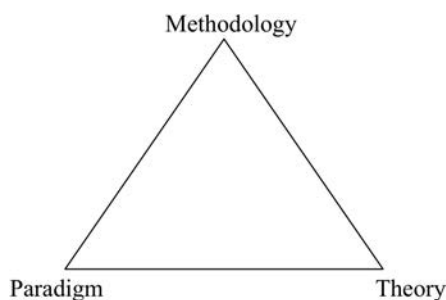


Figure 29.2 Aligning methodology with theory and paradigm

ontology) and the theory adopted. In other words, the researcher's stance on what constitutes reality should be consistent with the way it is investigated.

Understanding and Establishing Rigor in Qualitative Research

As noted, the Gurzynski-Weiss and Baralt (2014) study foregrounds the ontological and epistemological divide between ISLA researchers who operate from a postpositivist and a postmodernist perspective, respectively, which are mutually exclusive (Zuengler & Miller, 2006). While the former group (e.g., Gurzynski-Weiss & Baralt, 2014) strives for validity and reliability in their work, the latter group (e.g., De Costa, 2015 discussed earlier) seeks to establish trustworthiness and dependability of data, which is made possible through longitudinal study and a method of reporting known as *thick description* (i.e., using multiple perspectives to explain the findings from the study and the participants' insights). Admittedly, qualitative ISLA researchers working from a postmodern stance have come under attack for the subjectiveness of their work. In their defense, these researchers (e.g., Flyvbjerg, 2011) would argue that all research is subjective and is therefore susceptible to researcher bias, which may result in some findings being suppressed by the researcher and being analyzed from a skewed perspective. In short, an understanding of how rigor is interpreted and achieved along paradigmatic lines is vitally important because it helps explain why qualitative ISLA researchers (1) emphasize the importance of aligning their methodology and chosen SLA theory, and (2) insist on making transparent their working process (Holliday, 2015).

Articulating the Discourse Analytic Approach

Admittedly, qualitative researchers are not expected to articulate their paradigmatic stance explicitly in their work; nevertheless, this perspective is often easily inferred from the methodology and theoretical framework they adopt and the research questions that guide their studies. Researchers are, however, increasingly expected to describe their discourse analytic method, and explain how they coded and subsequently transcribed their data in ways that also align with the theoretical framework and methodology that guide their study. Fortunately, the availability of discourse analysis handbooks (e.g., Gee & Handford, 2013) and an increasing discussion of different approaches to discourse analysis within SLA have facilitated this possibility. In fact, the importance of describing how one's data are analyzed is underscored by the discussion and inclusion of main approaches to discourse analysis in the 2016 *TESOL Quarterly* research guidelines (Mahboob et al., 2016) mentioned earlier. Simply put, in addition to illustrating how the data were analyzed (i.e., discourse analytic approach), ISLA researchers are increasingly also expected to go beyond the content and face value of the data (i.e., thematic analysis), to explore, for example, how data collected were influenced by other things such as researcher positioning.

Equally important is the expectation that transcription conventions be included at the back of the study; as observed by Swann (2010), "transcriptions correspond to a researcher's interests and what they see as the analytical potential of their data, as well as their wider beliefs and values" (p. 163). Thus, how data is transcribed is not perceived as a neutral process; rather, transcription choices are often viewed as yet another indicator of the ISLA researcher's paradigmatic orientation (see Green, Franquiz, & Dixon, 1997 on the politics of transcription). For example, nonverbal aspects of communication can become the focus of the researchers who work in conversation analysis. Hence, a conversation analytic-oriented

transcription would reflect and capture these paralinguistic features because such features are central to the analysis of data (see Smotrova & Lantolf, 2013).

In his ethnographic study of the impact of emotions on language learning, De Costa (2016a), for example, described how he coded his data using Corbin and Strauss's (2008) coding system and analyzed the data by using ethnographic microanalysis of interaction (Garcez, 2008). Beginning with *open coding*, De Costa started by writing down anything that came to his mind. This initial step helped him bracket any preconceived assumptions while he looked for ways in which the data related to the emotions of his participants. This exercise was necessary given his familiarity with the research site. Next, he used the *axial and selective coding* processes of breaking down, examining, and conceptualizing his data. Axial coding (i.e., coding the central phenomenon) allowed him to assess whether the codes needed to be identified as categories, collapsed into other codes, or further separated into subcodes. Then, at the selective stage, he revisited the data that were organized into central categories, checked for data saturation, and searched for discrepant cases (for a detailed discussion on how to code SLA data, see also Baralt, 2012).

Exploring Researcher Reflexivity and Ethics

As mentioned at the outset of this chapter, the reflexivity or “presence” of the researchers in the accounts they present is pivotal in qualitative research (Creswell, 2012). This necessity is primarily because the qualitative researcher is often implicated in the research process, and therefore needs to clearly explain her individual history and her relationship to the research site (the classroom, in the case of ISLA) and her research participants, as they may have influenced the findings and analyses. Within SLA, Kramsch and Whiteside (2007) have called for researcher positioning “to be explicitly and systematically accounted for and placed in its historical, political, and symbolic context” (p. 918). In light of the strong influence that the researcher may potentially wield on the data collection and analysis process, qualitative methods such as interviews have come under scrutiny in recent years. In addition to being an instrument for data collection, Talmy (2010), for example, called for a rethinking of interviews as a social practice between the researcher and her participant, underlining how the two parties position each other inevitably shapes the general architecture and findings of an interview. Relatedly, the complex relationship between the researcher and the researched is increasingly being examined for any ethical infringements. Of growing interest is how researchers negotiate ethical dilemmas that emerge during the research process as part of the broader reflexive turn in applied linguistics (De Costa, 2016b).

Mixing Methods, Triangulation, and Aligning Theories With Different Methodologies

ISLA has witnessed a growth in the number of studies that combine methods (King & Mackey, 2016; Mackey & Gass, 2015). More studies now combine quantitative and qualitative research methods, because each highlights “reality” in a different yet complementary way (see Table 29.1). Mixed methods research (MMR), however, as Riazi and Candlin (2014) remind us, is highly complex and needs to be dealt with at the level of paradigms (see Figure 29.2), that is, whether the research is embedded within postpositivism or postmodernism. Also important to note is the distinction between MMR and triangulation. According to Phakiti and Paltridge (2015), while MMR involves the use of *both* qualitative and quantitative methods, triangulation “refers only to the strategy of collecting information from different or multiple sources to help gain a deeper understanding of a particular matter” (p. 15). To illustrate this

difference, they highlight how ethnographers who use a combination of interviews, observations and document analysis to answer their research questions are engaged in data triangulation but are not necessarily employing a mixed methods design. Drawing on Flick (2014) and Phakiti and Paltridge (2015) describe three common MMR designs:

- One-after-the-other design (qualitative data are added to counterbalance quantitative findings);
- Dominance design (more emphasis is given to one method over other methods);
- Side-by-side design (both methods are evenly balanced and are carried out concurrently to address different research questions).

Unfortunately, absent in most methodology sections in the ISLA literature is a full description of the extent of the role each method plays within the research design. In other words, currently lacking in most methodology descriptions are specificity of how qualitative methods are combined with quantitative methods.

As mentioned, there is no hard and fast rule that prohibits a blending of methodologies. Also noted earlier was how De Costa (2015) combined ethnography and case study in his ethnographic case study in a Singapore school. By the same token and in the spirit of innovation (Riazi, 2016), it is equally possible for ISLA researchers to investigate a SLA theory with different methodologies. In short, there is much potential and flexibility for ISLA qualitative research to be complemented by quantitative research and enriched by the growing number of SLA theories, which themselves have looked outward toward adjacent disciplines such as anthropology, sociology, and psychology for theoretical guidance. Put differently, it is possible for ISLA researchers to mix methods in ways that align with their chosen view of reality, which may be combined with a range of SLA theories. To ensure research transparency, the ISLA researcher would need to (1) provide a clear rationale for her aligning methodological and theoretical choices, (2) provide an account of how she effectively carried out her hybridized approach, and (3) discuss the type of MMR design adopted in her study.

Key Concepts

Ontology: A particular view of reality, which can be conceived, for example, as one and absolute (positivism), or multiple and co-constructed (constructivism).

Epistemology: In relation to one's view of reality, epistemology is the stance (objective for positivists or subjective for constructionists) on what constitutes knowledge.

Mixed Methods Research (MMR): An approach to research that combines or mixes different methods (qualitative and quantitative with dominance, concurrent or sequential designs) to render a multidimensional and more holistic view of the process under investigation.

Dominance design: A type of MMR in which more emphasis is given to one method over other methods.

Thick description: Detailed accounts of data in context with the aim of revealing the underlying structure and process of a phenomenon.

Discourse analytic approach: Methods to analyze language use and semiotic events according to the framework adopted. Examples of such approaches include conversation analysis, interactional sociolinguistics, genre analysis, narrative analysis, and critical discourse analysis.

Triangulation: Use of multiple data sources to provide a more comprehensive understanding of a certain phenomenon.

Peter I. De Costa et al.

Empirical Evidence

As noted, qualitative researchers are not expected to explicitly indicate their paradigmatic stance in their work because such a stance can generally be inferred from the research questions and their chosen methodology and theoretical framework. In this section, we exemplify five issues that were discussed in the previous section by drawing on sample studies. We also illustrate how various SLA theories and qualitative methodologies can be used to better understand the dynamics surrounding language learners.

Understanding and Establishing Rigor in Qualitative Research

Exemplar: Zappa-Holman and Duff (2015).

Research Concerns

This study sought to examine the socialization of a group of Mexican students into new academic cultures at a Canadian university by means of their social relationships, interactions, and other resources (e.g., material, symbolic) they accessed.

Primary Theoretical Framework

The study is conceived within language socialization theory (Duff, 1995; Schieffelin & Ochs, 1986), but it also draws on social network theory (Milroy, 1987) and community of practice theory (Lave & Wenger, 1991). Zappa-Holman and Duff developed the concept of individual network of practice (INoP) as a viable construct to examine how the students were socialized at university.

Methodology

It is a longitudinal, qualitative and multiple-case study that was conducted over a 12-month period.

Methods

Data sources were semi-structured interviews, writing logs, written materials, biographical, and academic data (e.g., years of prior English study, TOEFL scores), summary tables provided by the participants indicating the individuals in their respective social networks, and artifacts (e.g., course syllabi).

Participants

The study included 22 Mexican study abroad undergraduate students at the Canadian university. Of the 22 students, 6 were focal participants because of the richness of data they provided and their higher level of commitment to and interest in participating in the study. The article reports on examples drawn from three of the focal participants.

OUR COMMENTS

By focusing on three case participants and triangulating their rich and diverse data sets, Zappa-Holman and Duff were able to provide a thick description of the language socialization processes encountered by these students. This study is an example of how trustworthiness and dependability are constructed in qualitative research by making the researchers' working process transparent. Further, by literally mapping out the social networks their focal students participated in, the authors were also able to expand the SLA theoretical horizon by providing a dynamic alternative—their proposed concept of individual network of practice (INoP)—that focuses on the individual learner and her relation to her multiple contexts.

Articulating the Discourse Analytic Approach

Exemplar: Smotrova and Lantolf (2013).

Research Questions

(1) How is the gesture-speech unit enacted in the L2 instructional conversation that unfolds between teachers and students as they negotiate the meaning of new L2 concepts? (2) What evidence is there that the gesture-speech interaction between teachers and students mediates student understanding of L2 concepts?

Primary Theoretical Framework

This framework is based on Vygotsky's (1978) sociocultural theory.

Methodology

The authors explained how they carried out the videorecording and the ways in which the recordings (three selected excerpts) were then annotated using conventions associated with conversation analysis (ten Have, 2007). In line with the transcription conventions that were included as an appendix to their article, Smotrova and Lantolf also described how their transcriptions of simultaneous vocal and nonvocal actions were aligned, and how Russian utterances and English translations were presented in their data. Their transcribed data were also accompanied by still shots from the videorecordings in order to illustrate how their participants used gesture.

Methods

Data sources included 2 hours of videorecordings made in two separate EFL classrooms.

Participants

There were two nonnative EFL female instructors and two groups of undergraduate students in a tertiary institute in Ukraine. Both the instructors and students were Russian/Ukrainian bilinguals.

Peter I. De Costa et al.

Our Comments

Given that Smotrova and Lantolf sought to investigate how teachers and students deployed gesture-speech synchronization in order to mediate their understanding of L2 concepts, the authors selected an appropriately detailed discourse analytic approach—conversation analysis—that aligned with their research objectives of studying gesture. Indeed, conversation analysis enables researchers to closely examine how interlocutors build on each other's turns by using a range of semiotic resources. The study also illustrates the fruitful pairing of conversation analysis with an established SLA theory: Vygotskian sociocultural theory.

Exploring Researcher Reflexivity and Ethics

Exemplar: De Costa (2014).

Research Concerns

The study sought to investigate (1) the language ideologies embedded in the linguistic practices of five immigrant students and of the school they attended, (2) the discursive positionings of these students, and (3) how these discursive positionings and language ideologies impacted their language learning trajectories.

Primary Theoretical Framework

The frameworks of the study were language ideology (Kroskrity, 2010) and language identity (Norton, 2000). A more nuanced form of identity theory—positioning theory (Harré & van Langenhove, 1999)—was used in the larger study (cf., De Costa, 2012) to investigate how the participants were discursively positioned.

Methodology

This ethnographic case study appeared in the Research Issues section of *TESOL Quarterly* and thus has a strong methodological bent. De Costa focused on describing the ethical problems he encountered during a year-long study based in a Singapore secondary school. Before the study began, he secured institutional review board (IRB) approval and the support of the school principal and his participants. In reciprocation for their participation, he collaborated with his teacher participants by participating in lesson planning sessions and provided supplementary English lessons to his student participants. During the study, he avoided taking advantage of the teachers' hospitality by staggering his lesson observations. To minimize any student discomfort, only selected classroom interactions were videorecorded; the remaining interactions were audiorecorded. After the study, a deliberate effort was made to selectively disclose information shared with the principal to avoid harming his participants.

Methods

Data sources included structured interviews with students and teachers, classroom, school and excursion observations, and artifacts (e.g., samples of focal students' written work, student progress reports).

Participants

The participants were five focal immigrant students from China, Indonesia, and Vietnam who were 16 years old at the time of the study; eight Singapore students who were their classmates; and five of the focal students' teachers.

Our Comments

This study illustrates the messiness of conducting qualitative research in a nontertiary setting involving younger immigrant learners. As explained by De Costa, researchers need to exercise a level of flexibility while collecting data. Also emphasized is the need for the researcher to give back to his participants by positioning himself as a resource to them. Finally, ethical care, as noted by De Costa, also needs to be taken when presenting one's findings in order to ensure that no harm comes to the participants and that the findings are made accessible to multiple audiences.

Mixing Methods

Exemplar: Taguchi (2011).

Research Questions

(1) What patterns and pace of pragmatic development can we observe in the appropriate and fluent production of speech acts over one academic year? (2) Do individual differences and learning context affect the course of pragmatic development?

Primary Theoretical Framework

The theoretical framework is not clearly articulated, but it includes an implicit combination of Dynamic Systems Theory (de Bot, 2008), complexity theory (Larsen- Freeman & Cameron, 2008), and the emergentism approach (Ellis & Larsen-Freeman, 2006).

Methodology

Given that this was not an exclusively qualitative study, no qualitative methodology was identified in this 8-month longitudinal study. However, Taguchi did have a separate section where she described her role as a researcher and her relationship with the instructors in the EAP program in which her study was situated.

Methods

A 10-item survey was administered to document the student participants' amount of contact with English outside of class. In addition, a computerized oral discourse completion test was developed and administered three times over an academic year to quantitatively examine changes in pragmatic ability and thus address the first research question. Qualitative methods (interviews, observations, and journals) were used to examine how individual variation was related to the nature of their target language contact and experiences. The qualitative data were used to address the second research question.

Peter I. De Costa et al.

Participants

Forty-eight Japanese college students studying English in an English for Academic Purposes (EAP) program in Japan participated in the study. Twelve focal students were interviewed and the journal entries of 17 students were analyzed.

Our Comments

Given that both quantitative and qualitative methods in this study were evenly balanced and carried out concurrently to address the different research questions, Taguchi (2011) is an example of a mixed method study that bears a *side-by-side design* (Paltridge & Phakiti, 2015).

Unconventional Blending of Theories With Different Methodologies

Exemplar: Thompson and Vásquez (2015).

Research Question

How are the ideal and ought-to L2 selves expressed in the language learning narratives of highly successful language learners?

Primary Theoretical Framework

The framework of the study was L2 Motivational Self System (Dörnyei, 2009) and Self Discrepancy Theory (Higgins, 1987).

Methodology

The authors used narrative inquiry. They also provided a vivid account of how they transcribed, coded, and reviewed their data in order to identify their participants' paths to proficiency. They described their relationships with their participants and how they engaged in member-checking (i.e., checked to see if their participants agreed with the researchers' analyses) in order to preserve the ethical dimension of their research.

Methods

Data were collected through three in-depth narrative interviews. At the onset of each interview, participants were asked to discuss three main topics: (1) their earliest encounters with foreign languages; (2) how they decided to become teachers of a foreign language; and (3) any noteworthy experiences they encountered that related specifically to their status as non-native speakers of the languages they teach.

Participants

It is not clear how many teacher participants were part of the larger project that examined the lived experiences of nonnative foreign language teachers. However, in the article, Thompson and Vasquez focused on three teachers who taught three different languages: Italian, Chinese, and German.

Our Comments

Thompson and Vásquez's (2015) study is unique because the authors combined a narrative methodology with the L2 Motivational Self System (Dörnyei, 2009) and Self Discrepancy Theory (Higgins, 1987)—two SLA theories that traditionally have been investigated through the use of quantitative methodologies. In taking such an approach, the authors gained insights into learner motivation that may not have been accessible through the use of quantitative methods.

Future Directions

Over a decade ago, Lee and VanPatten (2003) underscored the need for language teachers to create language learning opportunities by tapping their SLA knowledge to reinforce acquisitional processes. How these opportunities for teaching and learning are investigated, however, has changed in the interim years, especially in light of the social turn in SLA (Block, 2003) and the fact that the traditional classroom as we know it has now expanded to include the virtual and global classroom. These changes, in turn, as we have argued in this chapter have prompted us to rethink our conceptualization of ISLA and how future qualitative ISLA research is to be carried out in order to better understand the acquisition dynamics surrounding L2 learners. Thus in terms of research settings, we anticipate that more research will be conducted in digital learning contexts with a view to examine how learning can be enhanced. One example is Lee (2014), who adopted Vygotskian socio-cultural theory to investigate how 15 advanced Spanish students used VoiceThread (an interactive multimedia tool) to create and exchange digital news regarding current events over the course of one semester. Her qualitative and quantitative data were gathered from multiple sources, including digital news recordings, reflections, online surveys, and final interviews. The study revealed that the creation of digital news stories in conjunction with a four-skills, integrated approach to task-based instruction promoted the development of learners' content knowledge and oral language development. In this vein, we expect that future research will continue to investigate both how digital affordances and Web 2.0 tools can be used to improve learning outcomes and how learners experience and react to these tools in the 21st century classroom.

In addition, and building on the growing body of study abroad research, we predict that more study abroad experiences will be examined qualitatively to rethink regular classroom instruction. Trentman (2013), who used an identity lens to investigate the language learning experiences of 54 students of Arabic on a study abroad program in Egypt, is a good case in point. Through a combined use of interviews, questionnaires, online technological observations, and participant observations, she illustrated how the degree of alignment between students' expectations and the realities they encountered in Egypt helped explain the extensive variation in the students' access to Egyptians and their use of the Arabic language. Pedagogically, insights from studies like Trentman's can be harnessed to manage learner expectations prior to a study abroad experience in order to optimize learning. Inevitably, venturing into unexplored teaching contexts—both online and abroad—will yield unprecedented ethical issues (Gao & Tao, 2016), which would need to be negotiated with care by reflexive ISLA researchers.

Further, and in order to facilitate qualitative research expansion into innovative configurations of the language classroom, “new” research methods would need to be added to the traditional basket of qualitative research tools. One tool that is increasingly used is focus

group interviews (Hennik, 2014). Focus groups, which have been used extensively as a tool for market research, typically consist of 6–8 participants who are preselected and have similar backgrounds. A focus group discussion is often led by a trained moderator who centers the discussion on a specific topic. In their investigation of the Vygotskian sociocultural influences on the use of a web-based tool for learning English vocabulary, Juffs and Friedline (2014) conducted focus group interviews to examine students' perceptions of classroom vocabulary learning vs. perceptions of vocabulary learning through the tutor. The interviews were moderated by an ESL teacher/researcher with whom the students were familiar. One key benefit of conducting focus group sessions is that they uncover a range of perspectives and experiences in a nonthreatening group environment.

Another qualitative research tool that can be developed for classroom research is open observation. To date, much of the observations in ISLA research have been closed in nature in that predefined categories are used for the observation schedule. One closed observation protocol that has been widely used is the Communicative Orientation of Language Teaching (COLT) observation scheme developed by Spada and Fröhlich (Spada & Fröhlich, 1995; see also Lightbown & Spada, 2013). However, to better capture the dynamics of a classroom and the emergent interactions that may occur among students and teachers, future qualitative researchers may wish to include open observations that comprise categories that emerge during the observation process. Such a shift to carrying out open observations would also be in line with the growing interest in grounded theory, an inductive methodology that is widely used in the social sciences (Charmaz, 2011). In contrast to the traditional model of research, where the researcher chooses an existing theoretical framework, and only then collects data to show how the theory does or does not apply to the phenomenon under study, grounded theory requires researchers to continually review their collected data, and group and regroup codes into concepts and categories, which then become the basis for new theory. To date, several SLA researchers (e.g., Back, 2011; Kubanyiova, 2012; Sato, 2013; Watzke, 2007) have adopted this methodology, and we anticipate that more researchers will use it in order to theorize and capture the fluid learning processes of L2 classrooms.

Throughout this chapter, we have also emphasized and demonstrated how qualitative and quantitative methods can be successfully paired. Building on the synergy that comes with bringing the two types of methods together, we predict that researchers will incorporate tools associated with corpus linguistics (Hyland, Chau, & Handford, 2012; Stubbs & Halbe, 2013), for example, into their investigative repertoire. One example is Yuldashev, Fernandez, and Thorne (2013) who examined L2 Spanish learners' multiword use with corpus linguistic tools and presented their data in relation to three case participants. This study is illustrative of how "big" data as derived from corpus research can be tapped and further supplemented by insights gleaned from case study research. In addition, and to take advantage of the affordances provided by technology, it is anticipated that more qualitative researchers will use coding software such as NVivo (Bazeley & Jackson, 2013) and programs such as Dedoose (<http://www.dedoose.com>), which is a program that is specifically designed for mixed-methods research and allows for the integration of text, audio, and video files (Silver & Lewins, 2014).

Finally, and in line with recent efforts to engage in theoretical and methodological cross-fertilization, we expect to see more hybridized research such as the recent work of Eskildsen and his colleagues (e.g., Eskildsen, 2014; Li, Eskildsen, & Cadierno, 2014), who have brought together conversation analysis and usage-based linguistics. Collectively, positive methodological developments coupled with growing theoretical diversity strongly suggest

that as a field, we have moved beyond a quantitative/qualitative and a cognitive/social divide that has long splintered SLA (DeKeyser, 2010; Zuengler & Miller, 2006). However, as we move forward, it is equally important that ISLA researchers not lose sight of the need to also study the traditions in which the varied methodologies are embedded and to find new ways to develop ground for further and future collaboration. At the end of the day, alignment at all levels—paradigmatic, theoretical, and ethical—needs to be taken into consideration.

References

- Atkinson, D. (Ed.). (2011). *Alternative approaches to second language acquisition*. New York: Routledge.
- Back, M. (2011). Legitimate peripheral participation and language learning: Two Quicha learners in a transnational community. *Language Learning*, 61(4), 1039–1057.
- Baralt, M. (2012). Coding qualitative data. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition* (pp. 222–244). Malden, MA: Wiley-Blackwell.
- Bazeley, P., & Jackson, K. (2013). *Qualitative data analysis with NVivo*. Thousand Oaks, CA: Sage.
- Benson, P. (2014). Narrative inquiry in applied linguistics research. *Annual Review of Applied Linguistics*, 34, 154–170.
- Benson, P., Chik, A., Gao, X., Huang, J., & Wang, W. (2009). Qualitative research in language teaching and language learning journals, 1997–2006. *Modern Language Journal*, 93(1), 79–90.
- Block, D. (2003). *The social turn in second language acquisition*. Edinburgh: Edinburgh University Press.
- Chapelle, C. A., & Duff, P. (Eds.). (2003). Some guidelines for conducting quantitative and qualitative research in TESOL. *TESOL Quarterly*, 37(1), 172–178.
- Charmaz, K. (2011). Grounded theory methods in social justice research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (4th ed., pp. 359–380). Thousand Oaks, CA: Sage.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). London: Sage.
- Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage.
- de Bot, K. (2008). Introduction: Second language development as a dynamic process. *Modern Language Journal*, 92(2), 166–178.
- De Costa, P. I. (2012). Constructing SLA differently: The value of ELF and language ideology in an ASEAN case study. *International Journal of Applied Linguistics*, 22(2), 205–224.
- De Costa, P. I. (2014). Making ethical decisions in an ethnographic study. *TESOL Quarterly*, 48(2), 413–422.
- De Costa, P. I. (2015). Re-envisioning language anxiety in the globalized classroom through a social imaginary lens. *Language Learning*, 65(3), 504–532.
- De Costa, P. I. (2016a). Scaling emotions and identification: Insights from a scholarship student. *Linguistics and Education*, 34, 22–32.
- De Costa, P. I. (Ed.). (2016b). *Ethics in applied linguistics research: Language researcher narratives*. New York: Routledge.
- DeKeyser, R. M. (2010). Where is our field going? *Modern Language Journal*, 94(4), 646–647.
- Denzin, N. K., & Lincoln, Y. S. (2011). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (4th ed., pp. 1–20). Thousand Oaks, CA: Sage.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Clevedon: Multilingual Matters.
- The Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, 100(S1), 19–47.

- Duff, P.A. (1995). An ethnography of communication in immersion classrooms in Hungary. *TESOL Quarterly*, 29(3), 505–537.
- Duff, P.A. (2014). Case study research on language learning and use. *Annual Review of Applied Linguistics*, 34, 233–255.
- Ellis, N., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics [Introduction to the special issue]. *Applied Linguistics*, 27(4), 558–589.
- Eskildsen, S.W. (2014). What's new? A usage-based classroom study of linguistic routines and creativity in L2 learning. *International Review of Applied Linguistics in Language Teaching*, 52(1), 1–30.
- Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal*, 81(3), 285–300.
- Flick, U. (2014). *An introduction to qualitative research* (5th ed.). London: Sage.
- Flyvbjerg, B. (2011). Case study. In N.K. Denzin & Y.S. Lincoln (Eds.), *The Sage handbook of qualitative research* (4th ed., pp. 301–316). Thousand Oaks, CA: Sage.
- Friedman, D. (2012). How to collect and analyze qualitative data. In A. Mackey & S.M. Gass (Eds.), *Research methods in second language acquisition* (pp. 180–200). Malden, MA: Wiley-Blackwell.
- Gao, X., & Tao, J. (2016). Ethical challenges in conducting text-based online applied linguistics research. In P.I. De Costa (Ed.), *Ethics in applied linguistics research: Language researcher narratives* (pp. 181–194). New York: Routledge.
- Garcez, P.M. (2008). Microethnography in the classroom. In K. King & N.H. Hornberger (Eds.), *Encyclopedia of language and education research: Methods in language and education* (Vol. 10, pp. 257–272). New York: Springer.
- Gee, J.P., & Handford, M. (Eds.). (2013). *The Routledge handbook of discourse analysis*. New York: Routledge.
- Green, J., Franquiz, M., & Dixon, C. (1997). The myth of the objective transcript: Transcribing as a situated act. *TESOL Quarterly*, 31(1), 172–176.
- Gullberg, M., & McCafferty, S.G. (Eds.). (2008). Gesture and SLA: Toward an integrated approach. *Studies in Second Language Acquisition*, 30(2), 133–146.
- Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in FTF and CMC modes. *Studies in Second Language Acquisition*, 36(1), 1–37.
- Harré, R., & van Langenhove, L. (1999). *Positioning theory*. Malden, MA: Blackwell.
- Hennik, M.M. (2014). *Focus group discussions: Understanding qualitative research*. Oxford: Oxford University Press.
- Higgins, E.T. (1987). Self—discrepancy: A theory relating self and affect. *Psychological Review*, 94, 319–340.
- Hinkel, E. (Ed.). (2005). *Handbook of research in second language teaching and learning* (Vol. 1). New York: Routledge.
- Hinkel, E. (Ed.). (2011). *Handbook of research in second language teaching and learning* (Vol. 2). New York: Routledge.
- Holliday, A. (2015). Qualitative research and analysis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 49–62). London: Bloomsbury.
- Hulstijn, J.H., Young, R.F., Ortega, L., Bigelow, M., DeKeyser, R., Ellis, N.C., Lantolf, J.P., Mackey, A., & Talmy, S. (2014). Bridging the gap: Cognitive and social approaches to research in second language learning and teaching. *Studies in Second Language Acquisition*, 36(3), 361–421.
- Hyland, K., Chau, M.H., & Handford, M. (Eds.). (2012). *Corpus applications in applied linguistics*. London: Bloomsbury.
- Juffs, A., & Friedline, B.E. (2014). Sociocultural influences on the use of a web-based tool for learning English vocabulary. *System*, 42, 48–59.
- Kasper, G., & Wagner, J. (2014). Conversation analysis in applied linguistics. *Annual Review of Applied Linguistics*, 34, 171–212.
- King, K.A., & Mackey, A. (2016). Research methodology in second language studies: Trends, concerns, and new directions. *The Modern Language Journal*, 100, 209–227.
- Kramsch, C., & Whiteside, A. (2007). Three fundamental concepts in second language acquisition and their relevance in multilingual contexts. *Modern Language Journal*, 91(1), 907–922.

- Kroskrity, P. (2010). Language ideologies: Evolving perspectives. In J. Jaspers, J. O. Östman, & J. Verschueren (Eds.), *Society and language use* (pp. 192–211). Amsterdam: John Benjamins.
- Kubanyiova, M. (2012). *Teacher development in action: Understanding language teachers' conceptual change*. Basingstoke: Palgrave Macmillan.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- Lazaraton, A. (2000). Current trends in research methodology and statistics in applied linguistics. *TESOL Quarterly*, 34(1), 175–181.
- Lee, J. F., & VanPatten, B. (2003). *Making communicative language teaching happen* (2nd ed.). New York: McGraw Hill.
- Lee, L. (2014). Digital news stories: Building language learners' content knowledge and speaking skills. *Foreign Language Annals*, 47(2), 338–356.
- Li, P., Eskildsen, S. W., & Cadierno, T. (2014). Tracing an L2 learner's motion constructions over time: A usage-based classroom investigation. *Modern Language Journal*, 98(2), 612–628.
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th ed.). Oxford: Oxford University Press.
- Lin, A. (2014). Critical discourse analysis in applied linguistics: A methodological review. *Annual Review of Applied Linguistics*, 34, 213–232.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
- Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design* (2nd ed.). New York: Routledge.
- Mahboob, A., Paltridge, B., Phakiti, A., Wagner, E., Starfield, S., Burns, A., . . . De Costa, P. I. (2016). *TESOL Quarterly* research guidelines. *TESOL Quarterly*, 50(1), 42–65.
- Markee, N., & Kasper, G. (2004). Classroom talks: An introduction. *Modern Language Journal*, 88(4), 491–500.
- Milroy, L. (1987). *Language and social networks* (2nd ed.). Oxford: Blackwell.
- Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change*. New York: Longman.
- Paltridge, B., & Phakiti, A. (Eds.). (2015). *Research methods in applied linguistics: A practical resource*. London: Bloomsbury.
- Phakiti, A., & Paltridge, B. (2015). Approaches and methods in applied linguistics research. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 5–25). London: Bloomsbury.
- Pica, T. (2012). Naturalistic and instructed language learning. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 455–458). New York: Routledge.
- Riazi, A. M. (2016). Innovative mixed-methods research: Moving beyond design technicalities to epistemological and methodological realizations. *Applied Linguistics*, 37(1), 33–29.
- Riazi, M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, 47(2), 135–173.
- Richards, K. (2003). *Qualitative inquiry in TESOL*. New York: Palgrave Macmillan.
- Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal*, 97(3), 611–633.
- Schieffelin, B. B., & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15, 163–191.
- Silver, C., & Lewins, A. (2014). *Using software in qualitative research*. Thousand Oaks, CA: Sage.
- Smotrova, T., & Lantolf, J. (2013). The Function of gesture in lexically focused L2 instructional conversations. *Modern Language Journal*, 97(2), 397–416.
- Spada, N., & Fröhlich, M. (1995). *Communicative orientation of language teaching observation scheme: Coding conventions and applications*. Sydney: NCELTR, Macquarie University.
- Spada, N., & Lightbown, P. (2012). Instructed second language acquisition. In P. Robinson (Ed.), *The Routledge encyclopedia of second language acquisition* (pp. 319–327). New York: Routledge.

- Starfield, S. (2015). Ethnographic research. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical resource* (pp. 137–152). London: Bloomsbury.
- Stubbs, M., & Halbe, D. (2013). Corpus linguistics: Overview. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–3). Hoboken, NJ: Wiley-Blackwell.
- Swann, J. (2010). Transcribing spoken interaction. In S. Hunston & D. Oakey (Eds.), *Introducing applied linguistics: Concepts and skills* (pp. 163–176). New York: Routledge.
- Taguchi, N. (2011). Pragmatic development as a dynamic, complex process: General patterns and case histories. *Modern Language Journal*, 95(4), 605–627.
- Talmy, S. (2010). Qualitative interviews in applied linguistics: From research instrument to social practice. *Annual Review of Applied Linguistics*, 30, 128–148.
- Talmy, S. (2011). The interview as collaborative achievement: Interaction, identity, and ideology in a speech event. *Applied Linguistics*, 32(1), 25–42.
- Talmy, S., & Richards, K. (Eds.). (2011). Qualitative interviews in applied linguistics: Discursive perspectives. *Applied Linguistics*, 32(1), 1–5.
- ten Have, P. (2007). *Doing conversation analysis: A practical guide*. Los Angeles: SAGE.
- Thompson, A. S., & Vásquez, C. (2015). Exploring motivational profiles through language learning narratives. *Modern Language Journal*, 99(1), 158–174.
- Trentman, E. (2013). Imagine communities and language learning during study abroad: Learning during study abroad: Arabic learners in Egypt. *Foreign Language Annals*, 46(4), 545–564.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Watzke, J. L. (2007). Foreign language pedagogical knowledge: Toward a developmental theory of beginning teacher practices. *Modern Language Journal*, 91(1), 63–82.
- Yuldashev, A., Fernandez, J., & Thorne, S. L. (2013). Second language learners' contiguous and discontiguous multi-word unit use over time. *Modern Language Journal*, 97(S1), 31–45.
- Zappa-Holman, S., & Duff, P. A. (2015). Academic English socialization through individual networks of practice. *TESOL Quarterly*, 49(2), 339–368.
- Zuengler, J., & Miller, E. (2006). Cognitive and sociocultural perspectives: Two parallel SLA worlds? *TESOL Quarterly*, 40(1), 35–58.
- Zuengler, J., & Mori, J. (Eds.). (2002). Microanalysis of classroom discourse: A critical consideration of method. *Applied Linguistics*, 23(3), 283–288.

Classroom-Based Research

Alison Mackey

Background

Many important discoveries within the field of second language acquisition (SLA) have emerged from carefully controlled studies of learners acquiring language in laboratory settings. While these studies offer important scientific advantages, much real-world language learning does not occur in laboratories, but in authentic contexts like instructed settings, in other words, second and foreign language classrooms. In order to better understand the relationship between instructional methods, materials, treatments, and second language learning outcomes, research needs to be carried out within the instructional settings where learning occurs. Instructed SLA research utilizes a full range of methodologies, ranging from the more quantitative to the more qualitative, as well as mixed and survey methods. For example, such methods include: (1) evaluating the effectiveness of teaching methods through experimental or quasi-experimental study designs; (2) analysis of teacher and learner behavior using observation protocols; (3) examination of interactional moves such as feedback sequences, negotiation of meaning, and language-related episodes (LREs) by recording, transcribing, coding, and analyzing segments of classroom discourse; (4) tapping into learner perspectives using introspective methods such as questionnaires, uptake sheets, learner diaries, interviews, and stimulated recall protocols; and (5) conducting ethnographic studies that strive for an emic view of the classroom.

In this chapter I first define and provide a brief overview of classroom-based language research by summarizing some of the most commonly used methods and procedures. Next I review some of the critical issues in second language classroom research methodology. Practical and logistical concerns and ways of addressing them when conducting research in classroom contexts are then discussed. Following this comes a description of some recent empirical work conducted in second language classrooms, where methodologies used by the researchers are highlighted. In the final section I explore future directions in this thriving and central domain of SLA research. Throughout the chapter I will highlight eight key concepts in the classroom-based research paradigm that provide quickly accessible information, intended to benefit experienced as well as novice researchers in the field.

Alison Mackey

Overview

Second language classroom research has been defined and described in a variety of ways over the years (e.g., Allwright, 1983; Bailey, 1999; Chaudron, 1988; Long, 1980; van Lier, 1990). As Allwright (1983) wrote:

Classroom-centered research is just that—research centered on the classroom, as distinct from, for example, research that concentrates on the inputs to the classroom . . . or the outputs from the classroom. It does not ignore in any way or try to devalue the importance of such inputs and outputs. It simply tries to investigate what happens inside the classroom when learners and teachers come together.

p. 191

In 2005, Nunan drew a distinction between classroom research and classroom-oriented research, saying

Classroom research includes empirical investigations carried out in language classrooms (however the term classroom might be defined). Classroom-oriented research, on the other hand is research carried out outside the classroom . . . but which make claims for the relevance of their outcomes for the classrooms.

p. 226

More recently, Williams (2014) has defined classroom research as “research in contexts with the following characteristics: the purpose is educational, an instructor is present, and more than one learner is present” (p. 541). As this brief overview shows, there are many definitions of classroom research. For the purposes of this chapter, I will adopt the definition by Gass and Mackey (2007) that classroom research involves “investigations carried out in second and foreign language classrooms, whether by the teachers of those classrooms or by external researchers” (p. 164).

Classroom-based studies are typically contrasted with research conducted in controlled laboratory contexts. Laboratory settings offer some distinct advantages, of course. In experimental settings, we are more easily able to assign learners randomly to treatment groups or control groups, to control or balance individual learner differences between groups, and to carefully control or mitigate other intervening variables. In the complex and often noisy domain of classroom-based SLA research, however, these variables are often impossible or impractical to control, and the placement of students in intact classrooms renders most classroom research quasi-experimental (i.e., nonrandom group assignment).

Key Concept

Quasi-experimental studies: Studies that lack random assignment to experimental or control groups as is the case in purely experimental studies. Many classroom-based studies are quasi-experimental because they utilize participants from intact classrooms. While nonrandom group assignment may be impossible to avoid when working in the classroom context, it may introduce confounding variables and therefore limit validity of the results.

Additionally, in classroom-based studies it can be difficult to create control groups without resulting in a potentially positive impact on learners in the experimental group(s) only which raises ethical issues and can result in challenges when seeking approval from institutional review boards (IRBs). Similarly, the goals of a classroom-based research project may not always align well with the plans or needs of the collaborating instructors. So classroom researchers often need to be flexible and responsive to the schools, administrators, instructors, and of course, the students they work with. Nevertheless, for the right research question, the benefits of conducting studies in authentic classroom contexts can outweigh these costs by providing a more authentic look at language learning processes and instruction (Hulstijn, 1997). Some authors have suggested that findings from laboratory settings cannot be readily applied to classroom settings at all. For example, Foster (1998), examining negotiation for meaning sequences in a classroom context, found results that contradicted those of experimental studies, and she proposed that study setting may significantly influence learner behavior. Whether or not findings from one setting can extend to the other is often an empirical question. However, clearly, findings from classroom research can complement those from laboratory-based studies, thereby painting a more complete picture of the complexities of second language learning (see Ellis, 1990, 1994; Norris & Ortega, 2000 for summaries of classroom research studies; also see McDonough & Mackey, 2013 for an overview of SLA research in diverse educational contexts).

Many have argued that research in a range of contexts, complemented with a variety of methods, is necessary to obtain a clearer picture of how languages are best learned and taught (e.g., Mackey & Gass, 2015). Traditional ideas about the purported dichotomy of quantitative, laboratory-based research versus more qualitative, classroom-based research is lessening and studies employing mixed methodologies are becoming more commonplace. A combination of both quantitative and qualitative techniques within a second language classroom research context adds to the methodological rigor of the investigation. Recently, some researchers (e.g., Hashemi & Babaii, 2013; Hulstijn et al., 2014; Ortega, 2005) have advocated for combining epistemological perspectives, and King and Mackey (2015) have argued against taking a narrow perspective when describing the importance of focusing on pressing real-world language problems, such as those that arise when working with language learners who have limited formal education, are residents or migrants from economically poor countries, are elderly, or are learning oral or literacy skills in an endangered language. Overall, conducting classroom-based research through a collaboration and combination of mixed and layered methods is a useful way of obtaining data on authentic instruction, interactions, language, and tasks that occur in second and foreign language classrooms.

Commonly Used Measures and Procedures

Rather than there being an exclusive set of methods utilized in classroom-based research, a variety of data collection techniques are implemented in classroom contexts much as they are in naturalistic, descriptive, or laboratory contexts. When applied in classroom contexts, these methodologies are given context-specific adjustments. Each method conveys benefits for data elicitation, as well as potential drawbacks. Space precludes a full discussion of all the techniques commonly used in second language classrooms, but some of the most frequently used are observation protocols, analysis of classroom discourse, questionnaires, uptake sheets, learner diaries, interviews, stimulated recall protocols, ethnographic observations, and experimental or quasi-experimental studies of the outcomes of different teaching methods. Among the most common are observations, introspective measures, and action

Alison Mackey

research techniques, each of which will be discussed here. I will also briefly present the research tools typically used in aptitude-treatment interaction (ATI) studies, which is an emerging but currently quite high-impact area.

Observations

Observations are a popular way to obtain comprehensive information about the events that actually occur in second or foreign language classrooms. Observations allow second language researchers to consider contextual variables occurring around instruction, tasks, peer-to-peer interactions, and student or teacher behaviors. The first step when planning out the procedures of a language classroom observation is to consider the goals of the research and to select an appropriate observation coding scheme to meet those research goals. Many coding schemes have been developed over the years (e.g., Allen, Fröhlich, & Spada, 1984; Allwright & Bailey, 1991; Chaudron, 1988; Fanselow, 1977; Guilloteaux & Dörnyei, 2008; Huang, 2011; Lynch, 1996; McDonough & McDonough, 1997; Mitchell, Parkinson, & Johnstone, 1981; Moskowitz, 1967, 1970; Nunan, 1989; Sinclair & Coulthard, 1975; Ullman & Geva, 1983) for a variety of different research questions and uses. Using existing observation protocols can be useful for new research studies if they appropriately apply to the research question at hand, but preexisting schemes are also helpful for researchers to look at when designing and devising their own observation protocols.

These coding schemes range from low-inference, meaning the coding of any easily observed behavior, to high-inference, including judgments about the function or meaning of a behavior. A low-inference scheme may simply be tallies of the number of questions asked over the course of a single class period (see Nunan, 1989). Two examples of high-inference schemes are the Target Language Observation Scheme (TALOS; Ullman & Geva, 1983) and the Communicative Orientation of Language Teaching (COLT; Allen et al., 1984; and more recently Guilloteaux & Dörnyei, 2008; Huang, 2011). These schemes are often adapted to address questions regarding the provision of corrective feedback. For descriptions of studies using these sorts of observational schemes see the section on Empirical Evidence.

Key Concept

Observation: A commonly used data collection method to obtain comprehensive information about the events that actually occur in second or foreign language classrooms. Common coding schemes utilized in observation studies include:

- The Target Language Observation Scheme (TALOS);
- The Communicative Orientation of Language Teaching (COLT);
- Corrective feedback coding schemes, among others.

There are many advantages to integrating an observation protocol into a classroom-based research study. First, it can help to uncover patterns of behavior and interaction that would be difficult to otherwise identify in a natural setting. Using an observation protocol ensures the systematic recording of relevant aspects of a classroom lesson and also facilitates coding the data. Furthermore, using or adapting a preexisting coding scheme can promote generalizability to previous studies and aid in the greater understanding of the inner workings of

language classrooms (Mackey & Gass, 2015). However, there are a few potential drawbacks to implementing observations. As with any observational study, predetermined categories of observation can act to limit researchers' perceptions of what occurred in the classroom, causing them to potentially miss important contextual features or other potentially relevant linguistic occurrences. Observations also limit the variety of data researchers have access to. For example, internal phenomena such as learner motivations, perceptions, and underlying cognitive processing are not readily observable and require other elicitation methods. One way to triangulate the rich data obtained from a classroom observation is to utilize introspective measures, the method we turn to next.

Introspective Measures

The following methods all involve the elicitation of learners' internal perspectives about their own language learning behaviors and experiences in the language classroom. In contrast to observations, where learning processes and outcomes are observed by a third-party researcher, introspective measures engage learners in data collection by encouraging them to communicate about the internal processes occurring during their language learning experiences. Access to this type of data is not possible from observational protocols alone. A variety of introspective measures, which vary in scope, have been implemented by SLA researchers. Three common types are uptake sheets, stimulated recall interviews, and diaries.

Key Concept

Introspective measures: A data collection method that involves the elicitation of learners' internal perspectives about their own language learning experiences in the language classroom. Access to this type of personal commentary is not possible from observational protocols alone. Three common types are:

- Uptake sheets;
- Stimulated recall interviews;
- Diaries/second language journals.

Uptake Sheets

Uptake sheets (see also Mackey & Gass, 2015) are a kind of worksheet designed for students to fill out during a given lesson or task, in which learners record what they notice about the language feature of interest or other aspects of the lesson or task. Originally developed by Allwright (1984, 1987), uptake sheets allow researchers to investigate learners' perceptions about what they are learning in real time and are useful for obtaining detailed, longitudinal data about classroom events. The researcher or instructor typically distributes uptake sheets at the beginning of a lesson and instructs learners to mark up the sheet as they proceed through the activity. Researchers have utilized uptake sheets to examine a variety of language classroom phenomena including learning processes, noticing of second language form, anxiety, and second language motivation. An uptake sheet that focuses on noticing, for example, may ask learners to record what they were noticing about different domains

Alison Mackey

of language (pronunciation, vocabulary, or grammar) during a classroom task, and then to report who produced the items they noticed (teachers, classmates, themselves), as well as if the item was new to the learner in that instance.

SLA researchers (such as Allwright, 1984; Almeida, 1995) suggest that uptake sheets can provide insights into the types of input learners attend to in the language classroom. Using uptake sheets, Nabei (2013) was able to compare what linguistic forms that 122 learners in a Japanese EFL reading class had noticed from their lessons with what forms the instructor had targeted in these lessons. The results of this study demonstrated that learners tended to focus mostly on vocabulary items and to a lesser extent on grammar, pronunciation, textual structure, and textual content, and that students' noticing of particular forms was related to instructor-led form-focused episodes in the classroom discourse. This noticing benefited the development of English vocabulary as well as reading strategies. Despite the variety of benefits of integrating uptake sheets in a classroom study, researchers should consider a few potential drawbacks. First, it is important to consider formatting and procedural decisions when designing and using uptake sheets. These decisions can affect the nature of what learners report, a potential threat to face validity. Learners should be carefully instructed to report on their own learning, rather than on what they think their teacher wants them to record. As with all introspective measures, timing is an additionally important consideration. If learners do not fill out the sheets until the end of class or after some period of time has elapsed, this could negatively impact the validity of the data, due to learners not being able to remember adequately what they noticed in class.

Stimulated Recall Interviews

Stimulated recall interviews (Gass & Mackey, 2016 provides a comprehensive overview of this methodology) have the goal of eliciting data about learners' thought processes at the original time of interactions, language tasks, and activities. Stimulated recalls prompt learners with a stimulus, such as an audio or video clip where they received corrective feedback during a task (Mackey, Gass, & McDonough, 2000), a sample of the writing they produced during class (De Silva & Graham, 2015), observational field notes (Do & Schallert, 2004), or a combination of these stimuli, and learners are asked to recall what they were thinking (or feeling) at the time. Stimulated recall has been utilized to examine cognitive and affective processes to better understand a variety of empirical questions such as second language strategy or inferencing use, second language teachers' decisions, second language writing choices and processes, second language reading and lexical use, and second language oral interaction, among other areas.

Key Concept

Stimulated recall interviews: Also known as verbal introspective reports, it is a technique that elicits data about the thought processes that take place during interactions, language tasks, and activities. Stimulated recalls prompt learners with a stimulus, such as an audio or video clip, before they are asked to recall what they were thinking and/or feeling at the time. This method can be useful for tapping into learners' motivations, thought processes, affects, noticing of linguistic input, or other data that would be otherwise unavailable to the researcher.

Some have questioned the extent to which verbal reports obtained through stimulated recall are valid and reliable (see Ericsson & Simon, 1993; Smagorinsky, 1994; and for empirical testing/treatment see Egi, 2007, 2008; Godfroid & Spino, 2015; Leow & Morgan-Short, 2004; Sanz, Lin, Lado, Wood Bowden, & Stafford, 2009; Smith, 2012). In Gass and Mackey (2016) we provide recommendations as to how stimulated recall methodology can be used to maximize potential benefits and mitigate the various pitfalls associated with this method. For example, we recommend that stimulated recall interviews should, wherever possible, be carried out as soon as possible following the event used as the stimulus. One of the potential pitfalls of this method is that the decay of short-term memory can cause learners to attempt to retrieve information from long-term memory, which is less reliable. For this reason, it is important that the stimulus for the recall is sufficiently strong so that learners are able to activate their memory structures. Audio and videorecordings work well but when more time passes between the initial event and the stimulated recall interview, more stimuli, such as transcripts, may be required. Additionally, in order to obtain data on what learners were thinking (and/or feeling) when they completed the task, rather than what they are feeling at the time of the interview, learners need to be properly trained. Therefore, pilot testing is essential when using a stimulated recall method. One way to train learners is by showing them a direct model and providing simple instructions. However, researchers should take care not to unnecessarily cue the learner or provide superfluous information. A final suggestion we make in Gass and Mackey (2016) is to consider allowing learners to control which aspects of the stimulus they wish to comment on. The less learners are led by researchers, the lower the chances of potential interference in the data are. However, structured interviews can lead to richer data that address researchers' questions. Therefore, these methodological decisions should always be made with the research questions in mind.

Diaries

Second language diaries, also known as second language or learner journals, are another introspective method useful for eliciting learner internal thoughts and processes while in the classroom. Learners or instructors are asked, with or without specific prompts, to write and reflect about their experiences learning or teaching the new language (see Bailey, 1983, 1990 for a detailed account of this method). Diaries written within classroom contexts can provide a range of useful data on a variety of aspects of the language learning process (Carson & Longhini, 2002; Oxford, Lavine, Hollaway, Gelkins, & Saleh, 1996; Rao & Liu, 2011). Important theoretical advances in SLA research have resulted or built on learner diaries, such as the well-known "Noticing Hypothesis" that emerged from Schmidt's diary (published in Schmidt & Frota, 1986) recording his experience learning Portuguese in Brazil. Some other areas that have been investigated with diary research include: foreign language anxiety (Bailey, 1983), issues of identity (e.g., Norton-Peirce, 1995), and second language motivation (e.g., Matsumoto, 1989). Diaries are also a good choice if researchers want to investigate contextual factors, such as the effects of a home stay or the influence of friends. Researchers may decide to give structure to the diary entries by providing prompts, or having the writers consistently journal about one aspect of a lesson over time. For example, the instructor or researcher may ask learners to write every day after peer work on the feedback they noticed giving or receiving from their peers.

Whether or not researchers ask learners to describe specific events in their learner journals via prompts, or simply ask learners to freely recount perceptions about their own language

Alison Mackey

learning experiences, second language diaries can be a useful addition to a classroom-based study. Like the other introspective measures described earlier, diaries tap into processes that would otherwise be difficult or impractical to elicit. Learners are often comfortable with this method as it is often used in general education classrooms as well as in language classrooms. Additionally, second language journals do not require specific data collection times and therefore are easier for learners or teachers to work into their daily routine. That being said, journal writing is also a commitment on the part of the learner that can be seen as burdensome. Diary data are also highly subjective; therefore, it is necessary to not over-generalize the perceptions of one particular group of language learners and their experiences to all learners.

Action Research

Also known as collaborative research, practitioner research, or teacher-initiated research, action research is often defined as “teachers doing research on their own teaching and the learning of their own students” (Crookes, 1993, p. 131). In action research, research questions emerge from a teacher’s own concerns and issues, rather than from theories deemed important by scholars in the research community. Thus, action research enables teachers to investigate topics unique to their own instructional situations and their own groups of language learners. Action research was discussed long ago by Lewin (1946) who outlined research steps instructors could engage in, including: (1) identify the problem, (2) carry out an action, (3) observe and reflect on the results, and (4) plan the next action (see Nunan, 1993, for a more comprehensive overview of the process involved in conducting action research). Atay (2008) identified some potential benefits of engaging in action research, such as the development of research skills, increased awareness of new teaching and learning practices and processes, and greater collaboration with colleagues and scholars in the field of education and language teaching. However, according to Wyatt (2011), despite the push for engagement in action research, instructors rarely do so unless pushed by professional development workshops or other teacher education.

Key Concept

Action research: A methodology that consists of investigations by practitioners on their own teaching and the learning of their own students. The steps for conducting an action research project include:

1. Identify the problem;
2. Carry out an action;
3. Observe and reflect on the results;
4. Plan the next action.

There are a variety of ways action research is implemented in practice. However, in any research plan the first step is always to identify the problem or concern that warrants further investigation. The identification of the main problem can be motivated by a teacher’s curiosity about things they see in the classroom, a desire to understand their classroom and students better, as well as any other professional development purposes. For example, a practitioner

might be concerned by the fact that students frequently misunderstand activity directions. A first step is to conduct a preliminary investigation; the practitioner in this case might gather information about why these misunderstandings are occurring in the classroom by observing their own instructions, and student behaviors and then examining problems that tend to occur. Preliminary data could be recorded as field notes, on an observational protocol, or through videorecorded class periods. Next, practitioners can take the data obtained and formulate a research question to test; for example, are there common vocabulary words in activity directions that students tend to misunderstand? Or, how can activity directions be improved to enhance student understanding? Then, instructors can design an intervention to test the effectiveness of changing their directions. Finally, there is a reflection stage that involves reexamination of the intervention.

While there are many benefits to engaging in action research, there are several drawbacks that instructors should consider before they embark on an action research plan. One drawback is that results tend to not be extendable to other contexts, because such research often responds to highly contextualized, local needs. Indeed, as others have pointed out (e.g., Nassaji, 2012), the knowledge gained through action research is not intended to be generalized. However, instructors might share the results of their research with colleagues who face similar challenges, groups of learners, and those who might wish to collaborate on future projects. Furthermore, action research often suffers from methodological limitations imposed by the real-world constraints of the classroom. For example, many action researchers are not able to utilize control groups for comparability purposes. Therefore, results should be carefully considered in terms of their validity and reliability. Mackey and Gass (2015) argue that if classroom action research is intended to be generalized and inform a wider community, it should meet the basic standards all studies are held to in the field of SLA research. That being said, it should be noted that there is no widely agreed upon criteria for evaluating the quality of action research (Burns, 2005).

Aptitude Treatment Interaction (ATI) Studies

ATI studies represent a growing trend in classroom research. Such studies seek to illuminate the relationship between the effectiveness of instructional treatments and the unique characteristics of individual learners by using a combination of classroom and experimental methodologies. ATI studies typically begin by investigating learner-internal individual differences in areas such as aptitude, working memory, cognitive creativity, motivation, learning styles, and learning strategies. Then, the effectiveness of instructional treatments (e.g., recasts, task sequencing) are measured and examined in light of these individual differences, either by comparison of group means or through correlational analyses.

A good example of an ATI study is an early study in this line of research, carried out by Révész in 2011. She examined the effects of task complexity and individual differences on form–meaning connections. This study is usually thought of as an ATI study because it investigates the relationship between the complexity of the task and learners’ affective individual differences to find out which version of the task is most effective for which types of learners. Participants were 43 ESL students from six different intact classes. These learners worked on two versions of the same argumentative decision-making task. The two versions, one complex involving higher levels of reasoning and more elements than the simple task, were recorded and coded for LREs, complexity, and accuracy of speech production. Additionally, questionnaires assessed the students’ individual differences: linguistic self-confidence, language use anxiety, and self-perceived communicative competence. An additional questionnaire

Alison Mackey

obtained information about the students' and teachers' perspectives and experiences with the task, including which version of the task was more useful for language learning, more difficult, more interesting, more stressful, more effective in drawing learners' attention to the quality of their output, and more successful in directing their attention to the quality of their peers' production. Results showed that in the complex task, learners were more accurate and showed more lexical diversity; however, their productions were less syntactically complex. The more complex task also induced more LREs. However, no effects of individual differences were found.

ATI research represents a relatively new area of inquiry that is still in the process of coming into fruition. SLA scholars (e.g., DeKeyser, 2009) have begun to call for more research of this type, as this approach to investigating aptitude-learning relationships holds great potential for illuminating how second language instruction can be optimized based on the unique needs of individual learners across a variety of pedagogical/instructional contexts.

Key Concept

Aptitude-treatment interaction (ATI): Studies that explore how learners' individual differences (e.g., aptitude, cognitive creativity, motivation, learning styles, strategies, working memory) are related to the effectiveness of varied kinds of instruction and pedagogical decisions. ATI studies empirically investigate how second language instruction can be optimized to fit the individual needs of a given learner.

Current Issues

In addition to choosing the appropriate study design and method for data collection for a classroom-based research study, there are many other issues, both practical and logistical, that should be considered. Some important current issues include the need to choose the best method and research design to use for assessing language development in the classroom context. There is an array of potential study designs and methods for measuring the effect of treatments commonly used in SLA research. For example, depending on the study aims, a cross-sectional design, a time series designs, or a pre-post design, with or without a delayed posttests and/or comparison or control groups may be used. In any case, the chosen method and study design should be appropriately matched to the research questions and context. Finally, there are several practical study design considerations that both external and internal classroom researchers should be aware of before embarking on a research project, which are discussed in the next section.

Measuring the Effect of Treatment

Classroom-based research has been instrumental in increasing our understanding of the importance of context and classrooms for instructed SLA. Many classroom-based studies have investigated the effect of instruction or interaction on development employing quasi-experimental designs. For example, studies examining the effectiveness of teaching pedagogies such as focus on form instruction (i.e., teaching grammatical features as they are

necessary for meaning-making as in task-based language teaching; see Long, 2015) and the effects of corrective feedback (see Mackey, 2012, for an overview) utilize outcome measurements to measure growth or development after a treatment.

One way for researchers to assess changes or development is to use a pretest/posttest design. The pretest serves to obtain baseline data from a group of learners before the treatment and comparability where possible. After a given treatment, participants take a posttest that is comparable to the pretest. The results from the posttest will allow researchers to examine the immediate effects of their treatment. For example, VanPatten and Cadierno (1993) examined the relationship between explicit instruction and input processing (i.e., perceiving the relationship between grammatical form and meaning) in a classroom-based study involving three intact classes as treatment groups. Pretest and posttest results were compared to determine which group outperformed the other two groups.

In order to measure longer term effects, such as retention of vocabulary items taught in class, a delayed posttest would be required. A delayed posttest should be comparable to the pretest and immediate posttest and administered sometime after the immediate posttest. Depending on the nature of the research question, multiple delayed posttests can be used several weeks or even months after the end of the treatment period. While delayed posttests are useful for examining how treatment effects change over time, the longer the delay the higher the likelihood of losing participants and the possible introduction of confounding variables such as maturation.

Key Concept

Delayed posttest: A developmental test that measures the longer-term effects of a given treatment. Delayed posttests are administered some time after an immediate posttest, such as 1 week later. However, multiple delayed posttests can be used 2 or even 3 months after treatment to better examine how treatment effects change over time.

For a variety of reasons, researchers may opt to implement a posttest-only design, where the focus of the study is primarily on performance rather than development. It may be necessary to use a posttest-only design if any pretest would give participants too much background information on what to expect from the treatment, or for other logistical reasons. In this case classroom researchers must take care to establish group comparability through other means such as a background questionnaire or another dependent variable like second language motivation or age of arrival. When comparability is a concern, researchers might also consider a repeated measures or a within-group design, in which each participant is assessed multiple times and their scores at different intervals are compared between the groups. In time-series designs, which are quite different, the amount of time allotted for pretests, observations, and posttests can vary by participant, allowing researchers to overcome comparability problems among their participants and avoid the use of a control group. As previously mentioned, it is often difficult to establish control groups in classroom-based research for logistical and practical reasons. Once the method of assessing development or performance and the design of the study have been chosen, classroom researchers should consider several practical concerns specific to classroom-based research next.

Alison Mackey

Practical Considerations

There are many practical concerns that need to be taken into consideration before embarking on a classroom-based study. With the exception of action research studies, classroom investigations are often conducted by researchers who are external to the context they are studying. For this reason, the presence of an external observer or data collector may influence results (see the often discussed Hawthorne Effect in the next section). Audio- and videorecording along with the distraction of having a new person in the classroom is also a challenge that researchers need to recognize, so that the research does not become obtrusive to learning. Also, classroom researchers need to ensure they have permission from an institutional review board (IRB), the school and program administration, the classroom teacher, students, and, if research involves children under 18, parents of students. Additionally, researchers have to secure permission to debrief all relevant parties on the results of the investigation. In situations, such as in action research, when the researcher is also the classroom instructor, issues of objectivity and subjectivity also need to be carefully considered. In general, classroom-based research requires ample flexibility, preparation, and patience and classroom researchers should take design choices, practical, and logistical issues into consideration prior to beginning their study. What follows is a more detailed overview of four of the most critical issues that every classroom-based study needs to consider.

The Hawthorne Effect

The Hawthorne Effect refers to the possibility that individuals who are being observed will modify their behaviors as a result of the observation. As most research methods textbooks note, the effect was first described by Brown (1954) and Mayo (1933) concerning observations that took place at the Hawthorne, Chicago, branch of the Western Electric Company. Workers at this company seemed to increase their productivity only when observers were present, leading the observers to be unable to capture an accurate picture of the working conditions.

Key Concept

The Hawthorne Effect: The possibility that individuals who are being observed will modify their behaviors as a result of the observation. First described by Brown (1954) and Mayo (1933), the Hawthorne effect can be minimized when classroom researchers utilize time-series designs.

The possibility of such an effect occurring in classrooms has led some to propose alternative designs to mitigate this concern. For example, Mellow, Reeder, and Forster (1996) argue that time-series designs are beneficial in this respect, as they involve many different instances of data collection both before and after treatment, over which participants gradually grow accustomed to being observed. Time-series designs are thus a useful method for reducing the Hawthorne Effect and allowing instructors and learners to become more natural during data collection.

Minimizing Disruptions

It is important to remember when conducting classroom-based research that it should not unduly disrupt the learning of the students or the teacher's instruction. During observations

of language classrooms, researchers should take several preliminary steps to ensure they do not disturb classroom activities, such as getting students accustomed to having the researcher or recording equipment in the room prior to beginning the study, ensuring observations do not conflict with the instructor's (or other observers') schedules, and asking the instructor for feedback in case they prefer the observer do something differently in subsequent observations. Debriefing the instructor during or after the observations (depending on the possible effects of disclosing information on the data) and thanking them for allowing the research are also key courtesies for maintaining a positive working relationship with an instructor and a school. Finally, researchers should take care to remember that their role in the classroom is not to judge or criticize (Murphy, 1992); therefore, they should always be sensitive to the perspectives of the instructor and students while the research is ongoing.

Maintaining Objectivity

While many classroom-based studies are carried out by external researchers, it is equally as likely that the researcher is familiar with the class, such as in the case of professional development observations, or even the instructor of the class under observation, as in the case of action research. For these reasons, issues of objectivity and subjectivity should be properly examined, and observers and researchers should be acutely aware how subjectivity could confound any of the variables in the study. For example, an instructor analyzing oral data collected from their own students might unknowingly evaluate them based on his/her preexisting knowledge of their students' abilities rather than from the data obtained alone. In cases where it is difficult for an action researcher to be objective about data, it can be useful to bring in an external coder for the data. Overall, issues of objectivity and subjectivity should always be carefully examined and accounted for at each stage of the study—data collection, data analysis, and interpretation of the results.

Institutional Review Board (IRB) and Informed Consent

Before any classroom study can begin, researchers must comply with a number of requirements from their home institutions as well as enlisting the support of all relevant personnel at the institution where the research is to be conducted. The first step is to obtain permission from both the IRB of the researcher's institution and any IRB at the program or school where the classroom is located. Obtaining IRB approval is often a long and arduous process, leading some to question whether IRBs do more harm than good (e.g., Schneider, 2015). Another important preliminary step is to obtain permission and enlist the support of the school administration and classroom teacher. Then it will be necessary to obtain informed consent from all relevant parties, typically meaning the instructor, students, and their parents (if they are under the age of 18). Informed consent documents usually are vetted by IRBs at both the researcher's institution and the participating school to ensure that none of the participants, instructors, or students feel coerced into participating in the study and that they know they can stop participating at any time for any reason. It is likely that over the course of a study some students will elect not to participate. In this case, every effort should be taken to accommodate these individuals, such as omitting any data that is inadvertently recorded or having those students sit behind any videorecording devices. In the case of all participants, every effort should be taken to maintain confidentiality of their data. Student data, such as grades or standardized test scores, represent highly sensitive information and the use of such data should be disclosed in all informed consent documents.

Empirical Evidence

The following section moves from methodological, design, and practical issues to consider before a study and describes recent empirical work in classroom contexts. These studies represent a range of classroom contexts, learners, designs, and methodologies. While some researchers utilize only one of the previously described methods (such as observational schemes) to obtain classroom data, others triangulate data elicited from several different methods. Results from these studies are presented to illustrate the authentic language instruction and development occurring in classrooms around the world.

Observational Studies

Several recent studies have utilized observation protocols to categorize classroom language data. A recent adaptation of the COLT observation protocol can be found in a study by Guilloteaux and Dörnyei (2008). In order to investigate the effects of teacher motivation strategy use on subsequent student motivation, the researchers utilized the real-time coding principles underlining the COLT scheme but changed the categories so that they measured motivational strategies (derived from Dörnyei, 2001). The new scheme was termed the motivation orientation of language teaching (MOLT). A self-report questionnaire triangulated student motivation data. A total of 27 language teachers and 1,381 students in 40 English classes at 19 different South Korean schools participated in the study. The researchers carefully defined and described the modifications they made to the existing COLT observation protocol to adapt it to their own research questions and listed the 25 observational variables that they deemed to measure teachers' motivational practices and learners' motivated behaviors. Relevant classroom events were recorded on the scheme every minute during observations. Results demonstrated that language teachers' motivational practices were in fact linked to learners' increased levels of motivated learning behaviors and their motivational states. Unlike previous motivation research that relied heavily on self-report survey data alone, the novel use of the MOLT observation protocol for observing motivational practices helped pave the way for further observational studies of motivation in language classrooms.

In one such study, Huang (2011) investigated the impact of content-based language instruction on young EFL learners' motivated behaviors and classroom verbal interactions by also using the COLT observation protocol. The researchers observed and recorded two instructors and 25 Taiwanese 6-year-olds from an intact primary school classroom four times during regular content or language-focused lessons. The author slightly adapted the COLT observation protocol to meet the unique needs of the study. The main difference, as in Guilloteaux and Dörnyei (2008), was the integration of Dörnyei's (2001) motivational system into the scheme so that student-motivated behaviors could also be recorded. In this way, the researcher was able to capture student attention, engagement, and amount of "eager volunteering." However, the other COLT categories remained the same as in the original COLT. The COLT observation data revealed that these young learners would participate more actively in content-focused rather than language-focused class periods. Learners were also recorded utilizing longer and more complex sentences in the content sessions as opposed to the language-focused sessions. In general, the use of the (adapted) COLT scheme allowed the researcher to easily compare the quality and quantity of the verbal interaction in the two types of classrooms. Additionally, the integration of materials from a previous, similar study (Guilloteaux & Dörnyei, 2008) allowed for comparability between the two studies.

Introspective Studies

In a classroom study, where we utilized the introspective method of uptake sheets, Mackey, McDonough, Fujii, and Tatsumi (2001) examined different methods of obtaining reports from learners and teachers on their perceptions of learning in an second language classroom. We wanted to know whether learners' reports were affected by the format of the uptake charts and the classroom context. Participants included 16 adult ESL learners in an intensive university English classroom. We compared three different uptake charts in the study, each with a similar format but a different focus. The varied foci included asking students what they noticed about pronunciation, language and context, or language structure (relative clauses, modals, etc.). In this study learners filled out the uptake charts during class time with the three varied foci rotated and counterbalanced. Results demonstrated that the format did in fact affect the quantity and quality of what learners reported during their language classes. The language focus format differed from the other two formats in terms of the amount of items it elicited (more than the other versions). The language and context format elicited more details on the specific items learners reported on. We concluded that careful design of uptake sheets is critical and should always include obtaining pilot data.

Some studies have sought to triangulate data obtained from multiple introspective methods. In Mackey (2006), I utilized data from journals and stimulated recall interviews to investigate the relationships between feedback, instructed ESL learners' noticing of second language form during oral interactions, and subsequent language development. This small-scale study involved 28 ESL learners from two intact speaking and listening university English classes and their two experienced ESL instructors. One class was randomly assigned as the treatment group that received interactional feedback, and the other class served as a control group that did not receive interactional feedback. I utilized online learning journals, stimulated recall interviews, and written questionnaires to measure noticing of second language form. The learning journals in this study were designed to elicit learners' impressions about interaction in their classroom as well as their impressions of the activities they completed and their overall learning during class. The learners filled out the journals three times a week for 4 weeks, and space was provided for learners to record which language forms they noticed (pronunciation, grammar, vocabulary, and content), who produced the forms they noticed (teacher, classmate, me, in the book), and whether the items were new to the learners or if the learner had heard the item before. Learners from the experimental group subsequently participated in stimulated recall interviews where I presented learners with videorecordings of 25 different feedback episodes from three different classroom activities. The learners were asked to report what they were thinking at the time of the feedback episode. Learners could comment on feedback that they received during the activities as well as feedback that their peers received. By triangulating results from both introspective methods, as well as pretest and posttest questionnaires, I found an interesting but complex relationship between interactional feedback in the classroom and the learners' reports about noticing of the feedback.

Another classroom study that integrated both observation and stimulated recall protocols was carried out by Bao, Egi, and Han (2011). The combination of both methods was used to investigate the extent to which uptake and stimulated recall can capture learners' noticing of recasts, an implicit form of corrective feedback that consists of a reformulation of the learner's erroneous utterance by an interlocutor. Twenty-five ESL student participants were first observed in their typical teacher-fronted classroom interactions and then immediately engaged in stimulated recall interviews that were coded for learners' noticing of feedback in

Alison Mackey

the form of recasts. Data were also coded for uptake following a recast—any response from the learners immediately following a recast, such as repeating or repairing their error. The authors included key contextual information in their reporting that was integral to understanding the results of their study. First, they carefully described the teaching styles of the instructors who participated in the study, comparing and contrasting how they interacted with their learners (one tended to provide feedback whole-class, while the other engaged in more one-on-one interactions). Second, the authors also detailed the extensive training they provided to the instructors where they introduced the concept of recasts and provided examples. The researchers also engaged the instructors in role plays where they modeled providing and responding to recasts. Using the data triangulated from both classroom observations and stimulated recall interviews, the researchers found that the rate of noticing recasts was higher when measured by stimulated recall data than by uptake measures. When learners mentioned they noticed recasts in their stimulated recall interviews, they were most often recasts accompanied by rising intonation.

Action Research

In order to promote the advancement of action research by language instructors, Wyatt (2011) reported on an inservice language teacher education course that included an action research component. Using observation and qualitative case study methodology, the author describes how four teachers engaged in action research as a result of a TESOL course at a Middle Eastern university. The author specifically focused on teachers' longitudinal development of action research skills over the course of 3 years and their growth in relation to using communicative tasks, designing materials, and developing literacy skills in their students. One instructor case study described Sarah, a high school English teacher who evaluated the effectiveness of communicative tasks she designed by audiorecording her lessons and through observation and field notes. By reflecting on the data she obtained in her action research, Sarah was able to identify points in her lessons where students engaged more in the communicative tasks as well as those tasks that led to greater acquisition of linguistic forms. Another instructor, Waleed, used action research to evaluate course materials he had designed and adapted. Waleed utilized observation protocols and interviews with fellow instructors to better understand how the course materials were being implemented in his school. His participation in action research led him to develop and teach professional development courses on how to utilize task materials creatively to support student motivation and learning within the classroom. Finally, Mariyam, a teacher trainer at the school, utilized stimulated recall interviews to help her fellow teachers reflect more critically on their own practice. She videorecorded her fellow teachers' lessons and engaged them in postlesson discussions about what they were thinking during the lesson and why. Mariyam used the results from these interviews to design workshops for the instructors. The author concludes that the action research utilizing multiple methods enabled the instructors in the case studies to achieve advanced research skills and improve their own instructional practices in addition to helping their peer instructors.

Calvert and Sheen (2015) conducted an action research study of task-based language teaching (TBLT) and learning that describes one teacher's experience with implementing tasks in her classroom. At the onset of the study, the instructor had no previous instruction in TBLT, but was teaching English for occupational purposes to 13 refugees and asylees in the US. Given the pedagogical challenges of instructing a group of learners with a variety of educational backgrounds, levels of English language proficiency, and time spent in the US, the instructor wished to integrate task-based teaching while documenting her reflections

and the challenges she overcame while implementing the program. In partnership with the second author, the instructor described how she designed her tasks and task evaluations and reported on the results. She additionally provided critical examinations and reflections on the results of each new task implementation, which helped the instructor to identify factors that posed barriers to effective task completion and to redesign these tasks. The instructor then implemented and evaluated the second iteration of the task and again critically examined and reflected on the results. The systematic evaluation of task implementation in her classroom allowed this instructor to obtain a greater understanding of her learners' needs and limitations and how she could best address those needs. The authors state that in conducting action research, the instructor learned how to modify a task-based pedagogical activity to improve its effectiveness and also learned about task-based teaching in general. The authors conclude by stating that the study "highlights the importance of action research as a means by which language teachers can address problems that arise in a TBLT lesson and, more generally develop their reflective skills" (Calvert & Sheen, 2015, p. 242).

Future Directions

There are many new trends in classroom research methodologies with researchers today aiming to incorporate multiple methods and layered approaches in their classroom-based research studies. As noted earlier, the aptitude-treatment interaction (ATI) studies are a relatively new development in this area with their focus on how learners' individual differences (e.g., aptitude, cognitive creativity, motivation, learning styles, strategies, working memory) are related to the effectiveness of varied kinds of instruction and pedagogical decisions. As noted, ATI studies empirically investigate how second language/FL instruction can be optimized to fit the individual needs of a given learner (see Goo, 2012; Li, 2015; Sheen, 2007; Yilmaz, 2013 for further examples of ATI studies) and they sometimes combine quantitative and qualitative methodologies and analyses, conducting some elements of the study in classrooms and other elements (such as working memory tests or other tests of individual differences) in the lab. This line of research has immediate and authentic implications for both language instructors and learners, as they consider the relationships among individual differences and the types of learning tasks they assign and practice in the classroom.

However, ATI research is just one of the many ways SLA researchers are utilizing classroom-based research to discover new insights into the processes of second language development. Without classroom-based SLA research, we would not have a more complete picture of how people learn languages in authentic situations. As this chapter has attempted to show, classroom researchers have a variety of methodologies, designs, and data elicitation techniques at their disposal that can be creatively combined to explore new questions concerning the teaching and learning of languages in authentic language learning contexts.

References

- Allen, P.B., Fröhlich, M., & Spada, N. (1984). The communicative orientation of language teaching: An observation scheme. In J. Handscombe, R.A. Orem, & B. Taylor (Eds.), *On TESOL '83: The question of control* (pp. 231–252). Washington, DC: TESOL.
- Allwright, D. (1983). Classroom-centered research on language teaching and learning: A brief historical overview. *TESOL Quarterly*, 17, 191–204.
- Allwright, D. (1984). Why don't learners learn what teachers teach? The interaction hypothesis. In D. Singleton & D. Little (Eds.), *Language learning in formal and informal contexts* (pp. 3–18). Dublin, Ireland: IRAAL.

Alison Mackey

- Allwright, D. (1987). Classroom observation: Problems and possibilities. In B. K. Das (Ed.), *Patterns of classroom interaction in Southeast Asia* (pp. 88–102). Singapore: SEAMEO Regional Language Center.
- Allwright, D., & Bailey, K. M. (1991). *Focus on the language classroom: An introduction to classroom research for language researchers*. Cambridge: Cambridge University Press.
- Atay, D. (2008). Teacher research for professional development. *ELT Journal*, 62, 139–147.
- Bailey, K. M. (1983). Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 67–103). Rowley, MA: Newbury House.
- Bailey, K. M. (1990). The use of diary studies in teacher education programmes. In J. Richards & D. Nunan (Eds.), *Second language teacher education* (pp. 215–226). Cambridge: Cambridge University Press.
- Bailey, K. M. (1999). *Looking back down the road: Twenty-five years of classroom research*. Retrieved August 1, 2005, from <http://www.linguistics.utah.edu/pdfdocs/Baileypaper.pdf>
- Bao, M., Egi, T., & Han, Y. (2011). Classroom study on noticing and recast features: Capturing learner noticing with uptake and stimulated recall. *System*, 39, 215–228.
- Brown, J. A. C. (1954). *The social psychology of industry*. Middlesex: Penguin.
- Burns, A. (2005). Action research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 241–256). Mahwah, NJ: Lawrence Erlbaum Associates.
- Calvert, M., & Sheen, Y. (2015). Task-based language learning and teaching: An action-research study. *Language Teaching Research*, 19, 226–244.
- Carson, J. G., & Longhini, A. (2002). Focusing on learning styles and strategies: A diary study in an immersion setting. *Language Learning*, 52, 401–438.
- Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.
- Crookes, G. (1993). Action research for second language instructors: Going beyond instructor research. *Applied Linguistics*, 14, 130–144.
- DeKeyser, R. (2009, October). *Variable interaction in SLA: Much more than a nuisance*. Plenary address given at the Second Language Research Forum, East Lansing, MI.
- De Silva, R., & Graham, S. (2015). The effects of strategy instruction on writing strategy use for students of different proficiency levels. *System*, 53, 47–59.
- Do, S. L., & Schallert, D. L. (2004). Emotions and classroom talk: Toward a model of the role of affect in students' experiences of classroom discussions. *Journal of Educational Psychology*, 96, 619–634.
- Dörnyei, Z. (2001). *Motivational strategies in the foreign language classroom*. Cambridge: Cambridge University Press.
- Egi, T. (2007). Recasts, learners' interpretations, and second language development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 249–267). Oxford: Oxford University Press.
- Egi, T. (2008). Investigating stimulated recall as a cognitive measure: Reactivity and verbal reports in SLA research methodology. *Language Awareness*, 17, 212–228.
- Ellis, R. (1990). *Instructed second language acquisition*. Cambridge, MA: Blackwell.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ericsson, K., & Simon, H. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Cambridge, MA: MIT Press.
- Fanselow, J. F. (1977). Beyond “Rashomon”—conceptualizing and describing the teaching act. *TESOL Quarterly*, 11, 17–39.
- Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19, 1–23.
- Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S. M., & Mackey, A. (2016). *Stimulated recall in second language research* (2nd ed.). New York: Taylor & Francis.

- Godfroid, A., & Spino, L.A. (2015). Reconceptualizing reactivity of think-alouds and eye tracking: Absence of evidence is not evidence of absence. *Language Learning*, 65, 896–928.
- Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven second language learning. *Studies in Second Language Acquisition*, 34, 445–474.
- Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of student motivation. *TESOL Quarterly*, 42, 55–77.
- Hashemi, M. R., & Babaii, E. (2013). Mixed methods research: Toward new research designs in applied linguistics. *Modern Language Journal*, 97, 828–852.
- Huang, K. M. (2011). Motivating lessons: A classroom-oriented investigation of the effects of content-based instruction on EFL young learners' motivated behaviours and classroom verbal interaction. *System*, 39, 186–201.
- Hulstijn, J. H. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition*, 19, 131–143.
- Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M., DeKeyser, R., Ellis, N. C., Lantolf, J. P., Mackey, A., & Talmy, S. (2014). Bridging the gap. *Studies in Second Language Acquisition*, 36, 361–421.
- King, K. A., & Mackey, A. (2015). Research methodology in second language studies: Trends, concerns, and new directions. *Modern Language Journal*, S(1), 209–227.
- Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26, 35–57.
- Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, 2, 34–46.
- Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, 36, 385–408.
- Long, M. (1980). Inside the “black box”: Methodological issues in research on language teaching and learning. *Language Learning*, 30, 1–42.
- Long, M. (2015). *Second language acquisition and task-based language teaching*. Hoboken, NJ: Wiley-Blackwell.
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. Cambridge: Cambridge University Press.
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27, 405–430.
- Mackey, A. (2012). *Input, interaction and corrective feedback in second language classrooms*. Oxford: Oxford University Press.
- Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design* (2nd ed.). London: Routledge.
- Mackey, A., Gass, S. M., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22, 471–497.
- Mackey, A., McDonough, K., Fujii, A., & Tatsumi, T. (2001). Investigating learners' reports about the second language classroom. *International Review of Applied Linguistics*, 39, 285–308.
- Matsumoto, K. (1989). An analysis of a Japanese ESL learner's diary: Factors involved in the second language learning process. *JALT Journal*, 11(2), 167–192.
- Mayo, E. (1933). *The human problems of an industrial civilization*. New York: Macmillan.
- McDonough, J., & McDonough, S. (1997). *Research methods for English language instructors*. London: Arnold.
- McDonough, K., & Mackey, A. (2013). *Second language interaction in diverse educational contexts*. Amsterdam: John Benjamins.
- Mellow, J. D., Reeder, K., & Forster, F. (1996). Using time-series research designs to investigate the effects of instruction on SLA. *Studies in Second Language Acquisition*, 18, 325–350.
- Mitchell, R., Parkinson, B., & Johnstone, R. (1981). The foreign language classroom: An observational study (*Stirling Educational Monographs*, No. 9). Stirling, Scotland: University of Stirling, Department of Education.

Alison Mackey

- Moskowitz, G. (1967). The FLint system: An observational tool for the foreign language class. In A. Simon & E. G. Boyer (Eds.), *Mirrors for behavior: An anthology of classroom observation instruments* (pp. 1–15). Philadelphia, PA: Center for the Study of Teaching, Temple University.
- Moskowitz, G. (1970). *The foreign language instructor interacts*. Minneapolis, MN: Association for Productive Teaching.
- Murphy, J. M. (1992). An etiquette for the nonsupervisory observation of second language classrooms. *Foreign Language Annals*, 25, 215–225.
- Nabei, T. (2013). Learner Uptake Reports on an EFL Reading Class in Japan. 関西大学外国語教育フォーラム, 12, 47–62.
- Nassaji, H. (2012). The relationship between SLA research and language pedagogy: Teachers' perspectives. *Language Teaching Research*, 16, 337–365.
- Norris, J., & Ortega, L. (2000). Effectiveness of second language instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Norton-Peirce, B. (1995). Social identity, investment, and language learning. *TESOL Quarterly*, 29, 9–31.
- Nunan, D. (1989). *Understanding language classrooms: A guide for instructor initiated action*. New York: Prentice-Hall.
- Nunan, D. (1993). Action research in language education. In J. Edge & K. Richards (Eds.), *Instructors develop instructors research: Papers on classroom research and instructor development* (pp. 39–50). Oxford: Heinemann.
- Nunan, D. (2005). Classroom research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 225–240). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ortega, L. (2005). For what and for whom is our research? The ethical as transformative lens in instructed SLA. *Modern Language Journal*, 89, 427–443.
- Oxford, R. L., Lavine, R. Z., Hollaway, M. E., Gelkins, G., & Saleh, A. (1996). Telling their stories: Language learners use diaries and recollective studies. In R. L. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 19–34). Manoa: University of Hawai'i Press.
- Palmeira, W. K. (1995). A study of uptake by learners of Hawaiian. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 127–161). Honolulu: University of Hawai'i.
- Rao, Z., & Liu, F. (2011). Effect of academic major on students' use of language learning strategies: A diary study in a Chinese context. *Language Learning Journal*, 39, 43–55.
- Révész, A. (2011). Task complexity, focus on second language constructions, and individual differences: A classroom-based study. *Modern Language Journal*, 95, 162–181.
- Sanz, C., Lin, H. J., Lado, B., Wood Bowden, H., & Stafford, C. A. (2009). Concurrent verbalizations, pedagogical conditions, and reactivity: Two CALL studies. *Language Learning*, 59, 33–71.
- Schmidt, R., & Frota, S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn* (pp. 237–326). Rowley, MA: Newbury House.
- Schneider, C. E. (2015). *The censor's hand: The misregulation of human subjects research*. Cambridge, MA: The MIT Press.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41, 255–283.
- Sinclair, J. M., & Coulthard, R. M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.
- Smagorinsky, P. (Ed.). (1994). *Speaking about writing: Reflections on research methodology* (Vol. 8). Thousand Oaks, CA: Sage.
- Smith, B. (2012). Eye tracking as a measure of noticing: A study of explicit recasts in SCMC. *Language Learning and Technology*, 16, 53–81.
- Ullman, R., & Geva, E. (1983). *Classroom observation in the second language setting: A dimension of program evaluation*. Toronto, Canada: Modern Language Centre, Institute for Studies in Education.

- van Lier, L. (1990). Classroom research in second language acquisition. *Annual Review of Applied Linguistics*, 10, 73–186.
- VanPatten, B., & Cadierno, T. (1993). Input processing and second language acquisition: A role for instruction. *Modern Language Journal*, 77, 45–56.
- Williams, J. (2014). Classroom research. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 541–554). London: Routledge.
- Wyatt, M. (2011). Teachers researching their own practice. *ELT Journal*, 65, 417–425.
- Yilmaz, Y. (2013). Relative effects of explicit and implicit feedback: The role of working memory capacity and language analytic ability. *Applied Linguistics*, 34, 344–368.

Experimental Research Methods

Kim McDonough

Background

As described in the introductory chapter to this volume (Loewen & Sato), instructed second language acquisition (ISLA) investigates how second language (L2) development is affected by systematic manipulations of learning mechanisms and conditions. In the current chapter, three experimental methods that have been used to explore topics in ISLA are introduced—structural priming, joint attention, and elicited imitation—all of which have explored the effect of manipulating learning mechanisms or conditions. These topics have been investigated predominantly through experiments carried out in laboratory settings, rather than in classroom contexts. In keeping with this section’s focus on research methods, the goal is to highlight the experimental techniques used to investigate these topics, rather than to provide a comprehensive history of each topic. For readers interested in more information about each topic, additional resources with more in-depth analysis have been provided.

Key Concepts

Structural priming: Facilitation in the processing of a structure due to previous experience with that structure.

Joint attention: The human capacity to coordinate attention with a social partner.

Elicited imitation: A testing technique in which a speaker is asked to repeat a series of sentences verbatim.

Structural Priming

Background

Structural priming is one type of repetition priming. Repetition priming refers to facilitation in the processing of language forms (phonological or structural) due to language users’ prior experiences with those forms. Within this category, structural priming specifically refers

to the tendency to produce a grammatical structure that appeared in the prior discourse as opposed to an alternate structure that could express similar message content (Bock, 1986). For example, if a speaker produces a relative clause, such as *it's the city that's got the Golden Gate Bridge*, later on she is more likely to produce another relative clause (*it's the museum that's featuring the Mona Lisa*) rather than an alternate structure, such as a prepositional phrase (*it's the museum with the Mona Lisa*) or a participle (*it's the museum displaying the Mona Lisa*). In order to elicit this phenomenon, structural priming experiments manipulate the linguistic forms present in the preceding discourse context in order to influence a speaker's subsequent language processing, production, or development.

Structural priming has been investigated through a variety of experimental methods dating from Bock's (1986) seminal work using a picture description task, in which participants heard and repeated prime sentences and then generated new utterances from keyword prompts to describe pictures. Additional experimental techniques for researching structural priming include oral and written sentence completion tasks, sentence recall, and scripted interaction (see the following Key Concepts box). Whereas the picture description, sentence completion, and sentence recall tasks typically involve an individual language user carrying out the task using a computer, the scripted interaction task requires communication between a participant and an interlocutor. Despite their differences, all four methods manipulate the form and order of primes and prompts in order to determine whether processing or production of a target structure is facilitated when that structure was present in the preceding discourse. For more information about structural priming, see overview articles written by first language (L1) researchers (Ferreira & Bock, 2006; Pickering & Ferreira, 2008) and methodologically oriented work by L2 researchers (McDonough & Trofimovich, 2008).

Key Concepts

Picture description: Along with distracter items, a participant hears sentences with the target structure and then uses key word prompts to generate sentences that describe pictures. The order of the sentence and picture trials is manipulated so that a prime sentence is heard before a target picture is described.

Sentence recall: Sentences are presented through rapid serial visual presentation, followed by a distracter task involving a word or number identification task. After completing the distracter task, a participant is asked to recall the preceding sentence.

Sentence completion: Sentence fragments are presented to participants who generate the rest of the sentence either orally (oral sentence completion) or in writing (written sentence completion). The amount of linguistic information provided in the prime fragments is manipulated to elicit a specific structure, but the target fragments can be completed using a variety of structures.

Scripted interaction: During conversation, an interlocutor who has been scripted with prime sentences interacts with a participant whose materials contain prompts. After the scripted interlocutor produces a prime sentence, the participant generates a new utterance using the prompts.

Current Issues

In addition to carrying out an extensive body of research to delineate the phenomenon of structural priming, researchers also have used structural priming as a methodological tool for investigating a variety of issues in psycholinguistics, L1 acquisition, and L2 processing and

development. Whereas early research focused on demonstrating that structural priming was a structural phenomenon that could not be attributed to the semantic, phonological, or lexical features of the prime and target sentences, more recent research has used structural priming techniques to describe the nature of children's mental representations (e.g., Rowland, Chang, Ambridge, Pine, & Lieven, 2012), such as to determine whether children have acquired abstract representations of target structures, as opposed to lexically specific representations including formulas or low-scope patterns. In addition, researchers have investigated how bilinguals store structural information (e.g., Bernolet, Hartsuiker, & Pickering, 2013), specifically whether they store grammatical information from each language separately or if the information is shared between languages. Another line of structural priming research has explored which aspects of structural priming can be attributed to a more implicit mechanism (such as its persistence over time) versus those that are more likely the result of an explicit mechanism (such as the effect of individual lexical items) (e.g., Kutta & Kaschak, 2012).

In terms of L2 speech production, researchers have investigated whether carrying out structural priming tasks facilitates subsequent production of target structures (e.g., Conroy & Antón-Méndez, 2015), and whether manipulations to the characteristics of priming tasks, such as the lexical features of primes and prompts (Kim & McDonough, 2008) or the explicitness of the target structure (Shin & Christianson, 2012), impact the occurrence and persistence of priming. Besides focusing on alternation between equally acceptable grammatical structures, such as active and passive constructions, L2 researchers have also investigated alternation between interlanguage and target language forms, including *wh*- questions (McDonough & De Vleeschauwer, 2012) and stranded prepositions (Conroy & Antón-Méndez, 2015). L2 researchers have also explored the occurrence of structural priming during peer interaction, during both face-to-face (McDonough & Chaikitmongkol, 2010; McDonough, Neumann, & Trofimovich, 2015) and synchronous computer-mediated communication (Collentine & Collentine, 2013) conversations. Corpus-based studies have also investigated whether structural priming occurs in naturalistic data as opposed to during experimental tasks (Collentine & Collentine, 2013; Thomas, 2016).

Empirical Evidence

Empirical evidence for structural priming is provided by calculating how frequently speakers produce a particular structure following exposure to that structure, compared to their use of that structure following exposure to an alternate structure. Structural priming generally is not concerned with the overall frequency of the two structures, but focuses on the association between the structure that speakers were exposed to and the structure of their response. It is expected that speakers will produce a structure more frequently following primes with the same structure than after primes with a different structure. For some studies that focused on L2 development, researchers provided only one structure in the prime sentences, typically the structure that they want the participants to produce. For example, rather than prime speakers to produce both active and passive sentences, researchers may target passive sentences only (Kim & McDonough, 2008), as they tend to be more difficult for L2 learners to produce. Similarly, if L2 speakers are alternating between target-like and interlanguage structures, researchers have primed them with the target-like structures only (Conroy & Antón-Méndez, 2015; McDonough & De Vleeschauwer, 2012). In these studies, empirical evidence for structural priming is demonstrated when speakers' produce target structures after the prime sentences more frequently than in contexts without a preceding prime sentence.

In terms of L2 research specifically, studies to date have demonstrated that structural priming occurs through experiments that tested a variety of English structures including passives (Kim & McDonough, 2008), phrasal verbs (Shin & Christianson, 2012), datives (Gries & Wulff, 2005; McDonough, 2006; Schoonbaert, Hartsuiker, & Pickering, 2007; Shin & Christianson, 2012), complex nouns (Bernolet, Hartsuiker, & Pickering, 2007), genitives (Bernolet et al., 2013), stranded prepositions (Conroy & Antón-Méndez, 2015), indirect questions (Biria, Ameri-Golestan, & Antón-Méndez, 2010), and relative and adverbial clauses (McDonough et al., 2015), along with Spanish nominal clauses (Collentine & Collentine, 2013) and French verb forms (Thomas, 2016). Similar to L1 research, L2 researchers have found that the occurrence of priming is affected by a number of factors. For example, the occurrence of shared lexical items in the primes and prompts, which is referred to as the lexical boost (Kim & McDonough, 2008), facilitates structural priming, although its effects may be more short term. In contrast, the number of intervening sentences between primes and prompts (Shin & Christianson, 2012) reduces the occurrence of structural priming. Although most studies have focused narrowly on students' production of target structures immediately following or one day after the priming activities, some L2 studies have found facilitative effects that persist for as long as 4–6 weeks (McDonough & Chaikitmongkol, 2010; McDonough & Mackey, 2008).

Future Directions

The L2 structural priming research to date has largely focused on whether priming occurs during L2 speech production by targeting a variety of structures and using different experimental techniques, such as picture description and scripted interaction. Fewer L2 studies, however, have investigated factors that have been shown to facilitate the occurrence of priming in L1 speech production, such as the semantic, homophone, and phonological boosts, plausibility, animacy, thematic roles, and participant role. Future research should investigate whether these factors also play a role in the occurrence or persistence of structural priming in L2 speech production. For L2 researchers interested in the contribution of implicit and explicit learning to L2 development, structural priming provides an ideal vehicle for manipulating factors that have been previously shown to increase its explicitness, such as the lexical boost, and comparing the longer-term effects of priming activities with varying levels of explicitness on speakers' subsequent use of the target structures.

For L2 researchers interested in issues such as cross-linguistic influence, the nature of the bilingual grammar, and L2 speakers' access to grammatical features, cross-linguistic structural priming shows promise as a methodological tool. Researchers are currently exploring whether the occurrence of cross-linguistic priming (i.e., providing primes in one language while eliciting responses in a different language) is contingent upon the structures having similar word order in both languages (e.g., Bernolet et al., 2007; Williams & Salamoura, 2007). In addition, studies about the effect of L2 proficiency on the occurrence of cross-linguistic influence have potential to shed light on the process by which L2 speakers' representations shift from being item-specific and language-specific to more abstract representations that are shared between languages (Bernolet et al., 2013).

Related to the issue of proficiency, an important topic for future research is whether L2 speakers must have mental representations of the target structure in order for priming to occur or whether priming can contribute to the formation of an initial representation (McDonough & Trofimovich, 2015). If L2 speakers have not formed abstract linguistic representations of the target structure, their production of those structures may be contingent on

Kim McDonough

their ability to reuse lexical items (such as repeated pronouns or nouns) or semantic features of the prime sentences. Consequently, in the absence of such features, priming of the underlying grammatical structure may be unlikely to occur. Recent structural priming research that targeted a completely novel construction (i.e., Esperanto transitives) found that L2 speakers' performance during a priming activity provided little evidence that new knowledge of the Esperanto transitive was acquired (McDonough & Trofimovich, 2015). Finally, when moving forward with structural priming studies, it may be useful to bear in mind Pickering and Ferreira's (2008) suggestion that "investigations *using* structural priming should not primarily be cast as investigations *about* structural priming" (p. 454, emphasis in original). In other words, using structural priming as a tool to investigate issues in applied linguistics and L2 pedagogy (such as task design and implementation) may be more useful than simply focusing on the occurrence of structural priming.

Joint Attention

Background

Joint attention refers to the human capacity to coordinate attention with a social partner. It occurs during conversation when interlocutors coordinate attention with each other by using and responding to visual cues such as gesture and eye gaze (Moore & Dunham, 1995). Using one's own eye gaze or gestures to lead an interlocutor to a common point of reference is known as *initiating* joint attention, while following the eye gaze or gesture of another person is *responding* to joint attention. A variety of visual cues can be used to initiate and respond to joint attention, such as head-turns, facial expressions, pointing, or eye gaze. Frequently used visual cues include interactive hand gestures (Bavelas, Chovil, Coates, & Roe, 1995), which serve to maintain interaction, such as seeking a response or coordinating turns, as opposed to convey lexical meaning, and conversational facial displays (Bavelas & Chovil, 1997), such as smiling and motor mimicry (e.g., wincing at another person's pain). Among visual cues, speaker eye gaze has the most consistent impact on listener responses (Bavelas, Coates, & Johnson, 2002). Joint attention is studied by exploring when speakers use such visual cues to initiate joint attention, and by identifying how their interlocutors respond to those visual cues. Evidence of listener's responses to speaker's visual cues include eye gaze, nodding, back channels, smiling, laughing, motor mimicry, gestures, supplying words or phrases, emotional displays, and dramatic intake of breath (Bavelas et al., 2002).

Key Concepts

Initiating joint attention: Using eye gaze or gestures to lead an interlocutor to a common point of reference.

Responding to joint attention: The ability to follow an interlocutor's visual cues, such as head-turns, pointing, or eye gaze.

Interactive hand gestures: Gestures that function to maintain interaction, such as by managing turn-taking, as opposed to gestures that communicate the meaning of words (such as pointing up when saying the word *above*).

Conversational facial displays: Facial expressions that react to an interlocutor's content, such as showing surprise, sympathy, or anger.

Current Issues

Early work about joint attention in developmental psychology (Scaife & Bruner, 1975) demonstrated that infants as young as two months can respond to their interlocutors' eye gaze, which stimulated questions about its potential role in helping children acquire language. For such young children, joint attention may help them segment the speech stream into individual words, associate an auditory form with its intended referent, figure out the meaning of individual words, and learn to combine words into utterances. Subsequent work with young children has reported high correlations between an infant's joint attention with caregivers and their language development (Carpenter, Nagell, & Tomasello, 1998; Morales et al., 2000; Tomasello & Farrar, 1986). Responses to joint attention are implicated in various forms of social and cognitive behaviours throughout life, including problem solving, mental and spatial rotation, visual scene processing, recognition memory, as well as language learning and use (e.g., Colonnese, Stams, Koster, & Nook, 2010; Dominey & Dodane, 2004; Kim & Mundy, 2012). Furthermore, failure to engage in joint attention is associated with learning deficits (Mundy, Gwaltney, & Henderson, 2010), and children with autism may have difficulty with the social function of joint attention (Jones, Carr, & Feeley, 2006). Compared to research in children's early language development, however, the role of joint attention in adult L2 learning is relatively underexplored. L2 studies with adults have focused more on gestures, such as pointing to referents, using motion to indicate directionality, and signalling locations, as opposed to eye gaze. This body of research has demonstrated that these methods of attracting joint attention may facilitate learning of L2 words (Gullberg, Roberts, & Dimroth, 2012; Kelly, McDevitt, & Esch, 2009; Macedonia & Knösche, 2011) and new sound contrasts (Hirata & Kelly, 2010; Kelly & Lee, 2012). For example, by pointing to an object, interlocutors can help learners make a deictic link between a sound string and its real-world referent, thereby helping facilitate form–meaning mappings. Similarly, by using gestures that convey information about the prosody and rhythm of speech (such as a hand flick or hand sweep), interlocutors can help learners perceive and produce sound contrasts.

Current work in L1 speech production has suggested that the eye gaze window may serve an important function in face-to-face conversation (Bavelas et al., 2002). Although mutual eye gaze between speakers and listeners often signals an exchange in roles (i.e., the listener becomes the speaker), Bavelas and colleagues identified a brief period of mutual eye gaze, which they termed the eye gaze window, in which speakers and listeners maintained eye gaze until the listener provided a verbal response, after which the normal gaze patterns resumed (i.e., listeners look at speakers more often than speakers look at listeners). In short, although speaker gaze to the listener initiates the gaze window, it is the listener's response that terminates the window. Recent L2 research has applied the eye gaze window findings to conversations between English L1 and L2 speakers, specifically whether eye gaze is associated with L2 speakers' responses to recasts (McDonough et al., 2015). They found that responses with a more target-like form were predicted by mutual eye gaze and the length of L2 speaker eye gaze during the response. The length of the L1 speaker's eye gaze to the participant while delivering the recasts was not predictive of target-like responses, which suggests that a successful feedback episode may be affected by interrelated speaker and listener gaze behavior (Goodwin, 1981).

Empirical Evidence

Evidence of joint attention is provided by identifying the visual cues that speakers' use to attract joint attention and documenting how listeners respond to those cues. Laboratory-based

Kim McDonough

research has shown that interlocutors spend a great deal of their interaction time looking at each other's faces (Argyle & Graham, 1976; Goodwin, 1981; Gullberg & Holmqvist, 2006; Kendon, 1967), with listeners tending to look at speakers for long looks with brief looks away, while speakers alternate looks to and away from listeners. It is believed that the speakers' looks to listeners put pressure on the listener to provide feedback or produce a response (Bavelas et al., 2002; Kendon, 1967). However, recent studies have shown that the typical eye gaze pattern is not maintained when speakers are asking questions (Rossano, Brown, & Levinson, 2009). Studies with young children have used structured video-taped assessments to assess joint attention, such as the Early Social Communication Scales (Mundy, Hogan, & Doehring, 1996), in which an experimenter and infant interact while seated facing each other at a table with various objects. Observations of their interactions result in two scores: the initiation of joint attention, which is based on how often the child uses eye gaze, pointing, or showing of objects to attract the attention of the tester, and responding to joint attention, which is the number of trials (out of six) in which the infant orients toward an object that the tester looks at and points to.

Empirical evidence for joint attention also is provided by coding gaze figurations based on videorecordings of the interactions (e.g., Rossano et al., 2009). Coding categories reflect a range of possible eye gaze configurations between speakers and listeners, ranging from the absence of any eye gaze toward the interlocutor to simultaneous eye gaze at their faces or eyes. In order to provide evidence that joint attention plays a role in language processing, researchers investigate the relationship between listeners' orientation to visual cues and their comprehension of the speakers' content. For example, Richardson and Dale (2005) reported that within zero to 6 seconds after a speaker has looked at an object, the listener also looked, with the listener's look most frequently occurring 2 seconds later. Furthermore, manipulating the listeners' orientation to visual information so that speaker and listener gaze was decoupled resulted in longer response latencies when answering comprehension questions. Studies to date, however, have not explored whether the same patterns are found with L2 speakers. For more information about technical considerations in using eye movement data for speech production research, see Griffin and Davison (2011).

Future Directions

Similar to structural priming research, joint attention studies also have potential to contribute to research about cross-linguistic influence. For example, cross-linguistic studies of L1 speakers' eye gaze has demonstrated that their orientation to visual information during face-to-face interaction is affected by structural differences in how languages express temporal-aspect domains (Von Stutterheim, Andermann, Carroll, Flecken & Schmiedtová, 2012). Whereas English speakers orient toward intermediate events in the sequence when watching video clips of motion events, German speakers pay more visual attention to the end point. A possible next step would be to examine how L2 learners of those languages orient to visual information, that is, as influenced by the L1 or the L2 structural patterns, and whether interaction with a more proficient interlocutor who uses visual cues to attract joint attention to the L2 visual orientation affects their language use. For a thorough overview of various hypotheses about the function of eye gaze during speech production, see Griffin (2004).

Joint attention also has potential to contribute to research situated within the view of conversation as a joint activity (Clark, 1996; Garrod & Pickering, 2009; Pickering & Garrod, 2004). In this framework, interlocutors establish successful communication by converging in their use of both linguistic forms (phonological, lexical, and grammatical structures) and

visual cues (posture, laughs, yawns). Structural priming studies have provided evidence that speakers align in their use of grammatical structures, such as by using a structure produced previously by an interlocutor. Just as alignment at one linguistic level can facilitate alignment at another linguistic level, convergence in visual cues, like eye gaze, may lead to shared linguistic representations. Research about joint attention during conversation (Richardson, Dale, & Kirkham, 2007) has shown that interlocutors use eye gaze to coordinate their attention toward pictures in an array, even when they are able to use other verbal back channels for directing attention. Subsequent studies have shown that eye gaze coordination is greater when interlocutors believe that visual information is not equally shared (Richardson et al., 2009). These studies, however, have focused on L1 speakers, so future studies might investigate how eye gaze facilitates convergence in L2 conversations, particularly whether eye gaze convergence is associated with linguistic convergence. In other words, studies might explore whether convergence in the eye gaze between interlocutors facilitates convergence in their use of linguistic forms.

Elicited Imitation

Background

Unlike structural priming and joint attention, both of which occur during real-world, natural language use situations, elicited imitation is testing technique that occurs in a highly artificial setting. As Vinther (2002) described it, the testing situation for elicited imitation is not natural as “no normal communicative situation requires the speaker to repeat series of sentences verbatim” (p. 54). The closest natural phenomenon to elicited imitation may be conversations between caregivers and children who are unable to repeat feedback, such as the “other one spoon” example documented by Braine (1971). Just as a child’s inability to repeat a caregiver’s utterance may be taken as evidence that the child’s linguistic system cannot accommodate the form, a speaker’s inability to repeat a stimulus sentence during an elicited imitation test is interpreted as revealing gaps in the speaker’s linguistic knowledge. In order for elicited imitation to provide insight into a learner’s linguistic system, as opposed to their repetition abilities or perceptual motor skills, it is crucial that the stimulus sentence be comprehended for meaning and regenerated using the speaker’s existing linguistic knowledge. Comparisons of the original and regenerated sentences are made to identify discrepancies that can provide insight into the speaker’s linguistic system.

Current Issues

Elicited imitation is currently being used in L2 acquisition research as a measure of implicit knowledge (e.g., Bowles, 2011; Ellis, 2005; Erlam, 2006; Spada, Shiu, & Tomita, 2015). Motivated by the goal of understanding the nature of L2 knowledge and the process by which it develops, researchers have worked to articulate key differences between implicit and explicit knowledge and to identify measurement tools that effectively assess both knowledge types. For example, Ellis (2005) proposed seven criteria that differentiate between explicit and implicit knowledge, such as awareness, self-report, systematicity, and examined how each type of knowledge related to a variety of assessment tools (elicited imitation, oral narrative, timed and untimed grammaticality judgment, metalinguistic knowledge). Elicited imitation, along with oral narrative and timed grammaticality judgment, were associated with implicit knowledge, whereas untimed grammaticality judgment and metalinguistic

Kim McDonough

knowledge measures were associated with explicit knowledge. Subsequent studies further validated the use of elicited imitation to measure implicit knowledge with other L2 learner groups (Bowles, 2011; Erlam, 2006; Spada et al., 2015) and used it to identify the effect of interventions on L2 development, such as feedback (Li, 2013) and form-focused instruction (Spada, Jessop, Tomita, Suzuki & Valeo, 2014). However, a recent study reported correlations between elicited imitation and metalinguistic knowledge tests (Suzuki & DeKeyser, 2015), which raises questions about the experimental conditions under which elicited imitation may capture implicit or explicit knowledge (or both).

Besides its use as a measure of implicit knowledge, currently elicited imitation is also used as a measure of oral L2 proficiency. In light of concerns that standardized measures of oral proficiency, such as the Oral Proficiency Interview, are expensive and time-consuming to administer and score, researchers have explored whether elicited imitation is a valid measure of oral proficiency (Cox, Brown, & Burdis, 2015; Tracy-Ventura, McManus, Norris, & Ortega, 2014; Wu & Ortega, 2013). This body of research has emphasized the need to include global proficiency measures in L2 acquisition studies, because traditional level descriptions such as “intermediate” or “advanced” and years of instruction can be difficult to interpret and compare across institutions. Elicited imitation can provide an independent measure of oral proficiency for inclusion in empirical research studies that is time efficient, economical, easy to administer, and easy to score. Several validation studies have shown that elicited imitation scores positively correlate with general indicators of proficiency, such as grades (Tracy-Ventura et al., 2014) and institutional levels (Wu & Ortega, 2013), as well as with oral proficiency interviews (Cox et al., 2015).

Empirical Evidence

For elicited imitation, empirical evidence of L2 speakers’ linguistic knowledge is inferred based on their ability to regenerate stimulus sentences, with their utterances assessed in terms of how accurately they reproduced the original sentences’ meaning or linguistic form. Stimulus sentences are usually typically presented aurally, but can also be presented visually. In order to minimize the possibility that participants imitate the sound sequences, as opposed to comprehend the meaning and regenerate the sentences, a short time delay (2–3 seconds) is often inserted between the presentation of the stimulus sentence and the cue for the participants to repeat. Alternatively, researchers may insert a cover task, such as answering a belief-statement comprehension question in order to delay repetition (Erlam, 2006; Suzuki & DeKeyser, 2015). The amount of time available for the speakers to articulate each sentence repetition may be limited, such as by providing only 8 seconds per response (Spada et al., 2015; Suzuki & DeKeyser, 2015), or response time may remain self-paced (Ellis, 2005; Erlam, 2006).

In order to ensure that the sentences are being regenerated, rather than imitated, their length is carefully controlled to ensure that it exceeds short term memory. The length of stimulus sentences in L2 research has ranged from 7–19 English syllables, 7–19 Chinese characters (Wu & Ortega, 2013), 7–19 French syllables (Tracy-Ventura et al., 2014), and 9–30 syllables for Russian (Cox et al., 2015). However, besides length, the lexical and syntactic features of stimulus sentence can also affect L2 speakers’ accuracy. Therefore, the frequency and familiarity of the vocabulary items used to construct the sentence can be controlled, such as by using words on frequency bands (Spada et al., 2015), and syntactic complexity can be addressed by controlling the number of clauses and morphemes (Tracy-Ventura et al., 2014). Although stimulus sentences have traditionally been grammatically accurate, researchers

have included both grammatical and ungrammatical sentences in order to elicit spontaneous corrections to the ungrammatical utterances (e.g., Erlam, 2006). The decision to include ungrammatical stimulus sentences has implications for the instructions. When all the stimulus sentences are grammatical, participants are generally instructed to repeat as much as they can or to repeat as well as they can. However, if some stimulus sentences are ungrammatical, then participants are informed to repeat in correct English what they hear, or are explicitly told to correct ungrammatical sentences (Suzuki & DeKeyser, 2015).

A final consideration in the use of elicited imitation to assess L2 speakers' linguistic knowledge concerns the scoring procedures. In child L1 research using elicited imitation tasks (commonly referred to as a sentence imitation), scoring is carried out based on error counts, often using an automated algorithm (e.g., Riches, 2012) that compares the original stimulus to the regenerated sentence and sums the number of words added, omitted, or substituted (including morphology). Alternatively, error counts are used as the basis of a categorical scoring system, such as the one outlined in the Clinical Evaluation of Language Fundamentals (Semel, Wiig, & Secord, 2003), where 3 points are awarded for verbatim recall, 2 points for one error, 1 point for two or three errors, and zero for more than three errors, with an error considered any deviation from the stimulus sentence (for an example study, see Poll et al., 2013). Categorical scoring has also been used in the oral proficiency L2 research; however, the criteria emphasized the semantic correspondence between the stimulus and regenerated sentences with less emphasis on deviations (Tracy-Ventura et al., 2014; Wu & Ortega, 2013). For example, whereas a verbatim repetition is awarded 4 points, a regenerated sentence with the same content but some grammatical or ungrammatical changes would be given 3 points. Furthermore, other L2 studies have scored elicitation imitation more narrowly by making a binary distinction based exclusively on whether the target structure was repeated accurately, with no consideration for the other elements in the sentences (Hirata-Edds, 2011; Li, 2013; Spada et al., 2015). For more information about the key methodological considerations for elicited imitation research, see Tomita, Suzuki, and Jessop (2009) and Vinther (2002).

Future Directions

As evidenced by the methodological differences in the design, administration, and scoring of elicitation imitation that were highlighted in the previous section, an obvious direction for future research is the validation of its methodological variants. The impact of methodological variation is particularly important for researchers who argue that elicited imitation is a measure of implicit knowledge, as changes to the nature of the stimulus sentences or the instructions may raise participants' metalinguistic awareness generally or their awareness of specific target structures (Chrabaszcz & Jiang, 2014; Spada et al., 2015; Suzuki & DeKeyser, 2015). Furthermore, scoring differences between the oral proficiency literature and the implicit knowledge research raise interesting questions about whether scoring differences reveal different aspects of L2 knowledge or proficiency.

Child language research to investigate specific language impairment has begun to focus on how performance on elicited imitation tasks reflects breakdowns in the various mechanisms involved in sentence regeneration, including the representation and retrieval of linguistic information in long term memory, maintenance of that information in short term memory and working memory, processing speed, receptive ability, and expressive phonology (Poll et al., 2013; Riches, 2012). Furthermore, this research has shown that the linguistic information in the stimulus sentences affects elicited imitation, ranging from lexical

Kim McDonough

features such as word frequency, abstractness, and imaginability, to prosody, syllable structure, semantic implausibility, intonation, and function words (Polišenská, Chiat, & Roy, 2015), with lexical and morpho-syntactic deficits in language knowledge, as opposed to memory, associated with poor performance. Future L2 studies might explore how a wide variety of processing abilities and learning mechanisms account for performance on elicited imitation tasks, in order to clarify the extent to which it assesses linguistic knowledge as opposed to other processing domains.

Conclusion

The use of structural priming, joint attention, and elicited imitation as experimental methods in ISLA has potential to provide further insight into how L2 learning is affected by systematic manipulations of learning mechanisms and conditions. Within the ISLA field, “acquisition” has been defined in many ways, with specific operationalizations typically reflecting theoretical perspectives about the nature of language and learning. Depending on how a researcher defines learning, the experimental methods discussed in this chapter can be adopted to provide insight into the conditions that facilitate or hinder the learning process. For example, structural priming experiments can shed light on L2 learning as operationalized in a variety of ways, such as the formation of an initial mental representation, strengthening of form–meaning connections, use during spontaneous production, decreased production of an interlanguage variant, or increased processing speed. Similarly, joint attention also allows researchers to focus on a specific point in the learning process, ranging from the initial identification of sound contrasts to pragmatic aspects of face-to-face communication.

Although each research method has its own origins, specific requirements, and logical applications, structural priming, joint attention, and elicited imitation can be exploited by L2 researchers to advance our understanding of key topics in applied linguistics research, ranging from the nature of linguistic knowledge to the social conditions that facilitate learning. Although these methods are primarily used in laboratory-settings to create greater experimental control, use in classroom settings could be possible with a few modifications. However, such modifications should be accompanied by validation information that helps clarify how changes to the task design, procedure, and coding may influence interpretations of task performance. Through careful adaptation of experimental methods to address specific topics of interest to L2 researchers, ISLA research can move forward without sacrificing its methodological rigor or ecological validity.

References

- Argyle, M., & Graham, J. A. (1976). The Central Europe experiment: Looking at persons and looking at things. *Journal of Environmental Psychology and Nonverbal Behavior*, 1, 6–16.
- Bavelas, J., & Chovil, N. (1997). Faces in dialogue. In J. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 334–346). Cambridge: Cambridge University Press.
- Bavelas, J., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, 21, 394–405.
- Bavelas, J., Coates, L., & Johnson, T. (2002). Listener responses as collaborative process: The role of gaze. *Journal of Communication*, 52, 566–580.
- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 931–949.

- Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition*, 127, 287–306.
- Biria, R., Ameri-Golestan, A., & Antón-Méndez, I. (2010). Syntactic priming effects between modalities: A study of indirect questions/requests among Persian English learners. *English Language Teaching*, 3, 111–119.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition*, 33, 247–271.
- Braine, M. (1971). On two types of models of the internalization of grammars. In D. Slobin (Ed.), *The ontogenesis of grammar: A theoretical symposium* (pp. 153–186). New York: Academic Press.
- Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4), 1–174.
- Chrabaszcz, A., & Jiang, N. (2014). The role of the native language in the use of the English non-generic definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, 30, 351–379.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Collentine, J., & Collentine, K. (2013). A corpus approach to studying structural convergence in task-based Spanish L2 interactions. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 167–187). Amsterdam, Netherlands: John Benjamins.
- Colonesi, C., Stams, G. J. M., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, 30, 352–366.
- Conroy, M., & Antón-Méndez, I. (2015). A preposition is something you can end a sentence with: Learning English stranded prepositions through structural priming. *Second Language Research*, 31, 211–237.
- Cox, T., Brown, J., & Burdis, J. (2015). Exploring proficiency-based vs. performance-based items with elicited imitation assessment. *Foreign Language Annals*, 48, 350–371.
- Dominey, P. F., & Dodane, C. (2004). Indeterminacy in language acquisition: The role of child directed speech and joint attention. *Journal of Neurolinguistics*, 17, 121–145.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172.
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27, 464–491.
- Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*, 21, 1011–1029.
- Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1, 292–304.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Gries, S., & Wulff, S. (2005). Do foreign language learners also have constructions? Evidence from priming, sorting and corpora. *Annual Review of Cognitive Linguistics*, 3, 182–200.
- Griffin, Z. (2004). Why look? Reasons for eye movements related to language production. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 213–247). New York: Psychology Press.
- Griffin, Z., & Davison, J. (2011). A technical introduction to using speakers' eye movements to study language. *Mental Lexicon*, 6, 53–82.
- Gullberg, M., & Holmqvist, K. (2006). What speakers do and what addressees look at: Visual attention to gestures in human interaction live and on video. *Pragmatics & Cognition*, 14, 53–82.
- Gullberg, M., Roberts, L., & Dimroth, C. (2012). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *IRAL*, 50, 239–276.

Kim McDonough

- Hirata, Y., & Kelly, S.D. (2010). Effects of lips and hands on auditory learning of second-language speech sounds. *Journal of Speech, Language, and Hearing Research*, 53, 298–310.
- Hirata-Edds, T. (2011). Influence of second language Cherokee immersion on children's development of past tense in their first language, English. *Language Learning*, 61, 700–733.
- Jones, A., Carr, E., & Feeley, K. (2006). Multiple effects of joint attention intervention for children with autism. *Behavior Modification*, 30, 782–834.
- Kelly, S.D., & Lee, A.L. (2012). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*, 27, 793–807.
- Kelly, S.D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, 24, 313–334.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22–63.
- Kim, K., & Mundy, P. (2012). Joint attention, social-cognition, and recognition memory in adults. *Frontiers in Human Neuroscience*, 6, 1–11.
- Kim, Y., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics*, 29, 149–154.
- Kutta, T., & Kaschak, M. (2012). Changes in task-extrinsic context do not affect the persistence of long-term cumulative structural priming. *Acta Psychologica*, 141, 408–414.
- Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal*, 97, 634–654.
- Macedonia, M., & Knösche, T.R. (2011). Body in mind: How gesture empower foreign language learning. *Mind, Brain, and Education*, 5, 196–211.
- McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition*, 28, 179–207.
- McDonough, K., & Chaikitmongkol, W. (2010). Collaborative syntactic priming activities and EFL learners' production of *wh*- questions. *Canadian Modern Language Review*, 66, 817–841.
- McDonough, K., Crowther, D., Kielstra, P., & Trofimovich, P. (2015). Exploring the potential role of eye gaze in eliciting English L2 speakers' responses to recasts. *Second Language Research*, 31, 563–575.
- McDonough, K., & De Vleeschauwer, J. (2012). Prompt type frequency, auditory pattern discrimination, and EFL learners' production of *wh*- questions. *Studies in Second Language Acquisition*, 34, 355–377.
- McDonough, K., & Mackey, A. (2008). Syntactic priming and ESL question development. *Studies in Second Language Acquisition*, 30, 31–47.
- McDonough, K., Neumann, H., & Trofimovich, P. (2015). Eliciting production of L2 target structures through collaborative priming activities. *Canadian Modern Language Review*, 71, 75–95.
- McDonough, K., & Trofimovich, P. (2008). *Using priming methods in second language research*. New York: Routledge.
- McDonough, K., & Trofimovich, P. (2015). Structural priming and the acquisition of novel form–meaning mappings. In S. Eskildsen & T. Cadierno (Eds.), *Usage-based perspectives on second language learning* (pp. 105–123). Berlin: Mouton De Gruyter.
- Moore, D., & Dunham, P.J. (1995). *Joint attention: Its origin and role in development*. Hillsdale, NJ: Lawrence Erlbaum.
- Morales, M., Mundy, P., Delgado, C., Yale, M., Messinger, D., Neal, R., & Schwartz, H. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, 21(3), 283–298.
- Mundy, P., Gwaltney, M., & Henderson, H. (2010). Self-referenced processing, neurodevelopment and joint attention in autism. *Autism*, 14, 408–429.
- Mundy, P., Hogan, A., & Doehring, P. (1996). *A preliminary manual for the abridged early social-communication scales*. Coral Gables, FL: University of Miami.

- Pickering, M., & Ferreira, V. (2008). Structural priming: A critical review. *Psychological Bulletin*, 134, 427–459.
- Pickering, M., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226.
- Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language Communication Disorders*, 50, 106–118.
- Poll, G., Miller, C., Mainela-Arnold, E., Donnelly Adams, K., Misra, M., & Park, J. S. (2013). Effects of children's working memory capacity and processing speed on their sentence imitation performance. *International Journal of Language Communication Disorders*, 48, 329–342.
- Richardson, D., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29, 1045–1060.
- Richardson, D., Dale, R., & Kirkham, N. (2007). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science*, 18, 407–413.
- Richardson, D., Dale, R., & Tomlinson, J. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science*, 33, 1468–1482.
- Riches, N. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language Communication Disorders*, 47, 499–510.
- Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. In J. Sidnell (Ed.), *Conversation analysis: Comparative perspectives* (pp. 187–249). Cambridge: Cambridge University Press.
- Rowland, C., Chang, F., Ambridge, B., Pine, J., & Lieven, E. (2012). The development of abstract syntax: Evidence from structural priming and the lexical boost. *Cognition*, 125, 49–63.
- Scaife, M., & Bruner, J. (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265–266.
- Schoonbaert, S., Hartsuiker, R., & Pickering, M. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, 56, 153–171.
- Semel, E., Wiig, E., & Secord, W. (2003). *Clinical evaluation of language fundamentals* (4th ed.). San Antonio, TX: Psychological Corporation.
- Shin, J., & Christianson, K. (2012). Structural priming and second language learning. *Language Learning*, 62, 931–964.
- Spada, N., Jessop, L., Tomita, Y., Suzuki, W., & Valeo, A. (2014). Isolated and integrated form-focused instruction: Effects on different types of L2 knowledge. *Language Teaching Research*, 18, 453–473.
- Spada, N., Shiu, J., & Tomita, Y. (2015). Validating an elicited imitation task as a measure of implicit knowledge: Comparisons with other validation studies. *Language Learning*, 65, 723–751.
- Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, 65, 860–895.
- Thomas, A. (2016). Syntactic priming in second language corpus data: Testing an experimental paradigm in corpus data. In A. Mackey & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research into second languages* (pp. 176–194). New York: Routledge.
- Tomasello, M., & Farrar, M. (1986). Joint attention and early language. *Child Development*, 57, 1454–1463.
- Tomita, Y., Suzuki, W., & Jessop, L. (2009). Elicited imitation: Toward valid procedures to measure implicit second language grammatical knowledge. *TESOL Quarterly*, 43(2), 345–350.
- Tracy-Ventura, N., McManus, K., Norris, J., & Ortega, L. (2014). Repeat as much as you can: Elicited imitation as a measure of oral proficiency in L2 French. In P. Leclercq, H. Hilton, & A. Edmonds (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 143–166). Bristol: Multilingual Matters.
- Vinther, H. (2002). Elicited imitation: A brief overview. *International Journal of Applied Linguistics*, 12, 54–73.

Kim McDonough

- Von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., & Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: Insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics*, 50, 833–867.
- Williams, J., & Salamoura, A. (2007). Processing verb argument structure across languages: Evidence for shared representations in the bilingual lexicon. *Applied Psycholinguistics*, 28, 627–660.
- Wu, S. L., & Ortega, L. (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals*, 46, 680–704.

Ethics in ISLA

Susan Gass and Scott Sterling

Ethics is knowing the difference between what you have the right to do and what is the right thing to do.

Potter Stewart, US Supreme Court Justice, 1958–1981

Background

Research into ethical practices in applied linguistics has seen increased activity over the last decade or so (De Costa, 2014; Fox, Artemeva, Darville, & Woods, 2006; Ngo, Bigelow, & Lee, 2014; Wen & Gao, 2007; Yeager-Woodhouse & Sivell, 2006), with publications discussing various aspects of what ethical research should look like for the field. Recent research related to ethics training has shown that applied linguistic and second language scholars do receive training in the topic but tend to do so through the guise of institutional review board (IRB) certification or informal training (Sterling, Winke, & Gass, 2016). Trust and opinions of science in general are low in various political and ideological factions (Hamilton, Hartter, & Saito, 2015), with new reports of research misconduct being reported daily (see <http://retractionwatch.com> for an updated list of scientific retractions). The fields of applied linguistic and second language (L2) research are not immune to retractions, with recent instances of plagiarism and data falsification being the source of retractions. Because of such mistrust and because ethical behavior must be the basis for any scientific inquiry, it is critical for research to be conducted and reported ethically to ensure that the larger L2 education world can trust the results put out by applied linguists. One way of combating any potential mistrust is to ensure that those conducting research are adequately trained in rigorous and ethical methodologies and are familiar with ethical practices. This chapter aims to shed light on a range of ethical issues that arise in research conducted in second language classrooms and to provide guidance as researchers navigate difficult ethical and moral decisions.

Susan Gass and Scott Sterling

Current Issues

Framing Research Ethics

Research ethics can be viewed from a wide range of perspectives. In its most general conceptualization, it has been defined by the British Economic and Social Research Council (ESRC, 2015, p. 43) as “the moral principles guiding research, from its inception through to completion and publication of results and beyond—for example, the curation of data and physical samples after the research has been published.” They outline six principles for ethical research conduct (p. 4):

- Research participants should take part voluntarily, free from any coercion or undue influence, and their rights, dignity and (when possible) autonomy should be respected and appropriately protected.
- Research should be worthwhile and provide value that outweighs any risk or harm. Researchers should aim to maximise the benefit of the research and minimise potential risk of harm to participants and researchers. All potential risk and harm should be mitigated by robust precautions.
- Research staff and participants should be given appropriate information about the purpose, methods and intended uses of the research, what their participation in the research entails and what risks and benefits, if any, are involved.
- Individual research participant and group preferences regarding anonymity should be respected and participant requirements concerning the confidential nature of information and personal data should be respected.
- Research should be designed, reviewed and undertaken to ensure recognised standards of integrity are met, and quality and transparency are assured.
- The independence of research should be clear, and any conflicts of interest or partiality should be explicit.

Hopkins (2014) points out (referring to an early version of the ESRC guidelines) that these principles are intentionally vague, allowing “researchers to make ethical decisions based on their own morality and the issues raised by their research project” (p. 72). This is an interesting interpretation because it implies that some decisions are not based on absolute ethical standards, but are open to interpretation. Duff and Early (1996) make a similar statement:

professional organizations and institutional review boards attempt to establish regulations to ensure that human research subjects are treated ethically. These principles and guidelines are undoubtedly helpful to researchers, but they may be neither self-evident nor absolute. Thus, interpretations or judgments often reside with the individual researcher or team.

p. 21

It is important to point out that research ethics training in applied linguistics and other research fields has tended to be viewed from the standpoint of procedural ethics (Guillemin & Gillam, 2004) or the types of research ethics needed to gain IRB approval; yet, social justification for research is also a significant issue. In fact, the six principles just listed encompass both procedural ethics and ethics related to the justification of research. The

early focus on procedural ethics can be seen in Schachter and Gass (1996, p. 173), where four points of consideration for ethical research that align with those of the ESRC are discussed:

- What are the official guidelines for the particular research project (research question, site, time frame)?
- Even if all official guidelines have been met, are subjects being treated fairly and with dignity (Are privacy and confidentiality ensured, and are the findings appropriately contextualized)?
- Are control treatments ethically correct?
- How are conflicts handled and reported?

Within the context of L2 research ethics, publications have appeared in SLA journals dating as far back as 1980 (Tarone, 1980). But, it was not until closer to the turn of the 21st century that the field saw a proliferation of writing on the topic (Cumming, 2002; Davies, 1997; Hafernik, Messerschmitt, & Vandrick, 2002; Norris & Ortega, 2000; Hamp-Lyons, 1998). Trends over the last 20 years have included a focus on experiential accounts of personal ethical issues (De Costa, 2014; Hobbs & Kubanyiova, 2008; Lee, 2011; Li, 2011), position papers (Ortega, 2005a; Ortega & Zyzik, 2008; Shohamy, 2004), and recently, quantitative analyses of various ethical issues (Sterling, 2015; Sterling et al., 2016).

Social Justification

In recent years there has been a turn in the field toward a more social view of research ethics (De Costa, 2014; Ortega, 2005a). Journal publications have begun to focus on topics such as the social utility of research (Ortega, 2005a) as well as reflective pieces concerned with various ethical issues that authors confront in their own research (Hobbs & Kubanyiova, 2008; Lee, 2011). In the case of the former, Ortega argues that all social science research “has as its ultimate goal the improvement of human life” (p. 430). Given this assumption, one must consider how research conducted in any social science field reflects that goal. She proposes three normative principles for research. The one most relevant for our discussion is that the value of research should be judged by and for its social utility. She further puts this in the context of instructed SLA research when she says (p. 430) “the value of instructed SLA research—just like the value of any other kind of social and educational research—ought to be judged not only by internal criteria of methodological rigor as understood by the particular epistemological models adopted, but also ultimately on the basis of its potential for positive impact on societal and educational problems.” Reflective pieces related to instructed SLA have dealt with topics such as the “challenges of engaging busy language teachers in one’s research, sustaining their commitment throughout the project and handling the physical and emotional strain of the researcher” (Hobbs & Kubanyiova, 2008, p. 495). Others (e.g., Koulouriotis, 2011; Lee, 2011) have dealt with the challenges of dealing with specific populations (e.g., nonnative speakers in the case of Koulouriotis) or ethical issues surrounding instances of racialization that would be revealed through research.

In this century alone, two journal special issues focusing on research ethics have been published in major SLA journals; issue 3 of volume 89 of the *Modern Language Journal* titled “Methodology, Epistemology, and Ethics in Instructed SLA Research” (Ortega, 2005b) and issue 5 of volume 28 of *TESL Canada Journal* (Kouritzin, 2011). The *TESL Canada Journal* special edition focused almost entirely on research ethics, but largely from

Susan Gass and Scott Sterling

the perspective of individual experiences, making it somewhat difficult to expand to a larger audience.

The articles in the *Modern Language Journal* edition focused on three large topics, methodology, epistemology, and ethics, with ethics receiving the least amount of coverage. In the introduction to the special issue, Magnan (journal editor, 2005) posed ethical questions, such as: Who is the audience of our research? What are the relationships among researcher, learner and teacher? What social, political, and human consequences, either intended or unanticipated may result from our investigations? (p. 315). Magnan (2005) went on to illuminate the ethical problems facing journal editors when she says:

Articles have been rejected for publication in professional journals, including the *MLJ*, because they were considered either so descriptive and local in nature that they would not inform a wide readership or so narrowly focused on how to learn or teach that they did not engage with theoretical issues. Might it be that, in these negative publication decisions, methodological and epistemological considerations weighed more heavily than the ethical concerns about the purpose and audience of our work?

p. 315

In addition to the special issues of journals that address research ethics, papers in Schachter and Gass (1996) chronicle the issues and challenges of conducting classroom-based research. Even though some of the challenges were not always framed as ethical issues, they clearly do relate to the ethical decisions that need to be made when one enters a classroom to conduct research. Duff and Early (1996) bring ethical issues to the forefront with their list of ethical considerations that include privacy/confidentiality, security (e.g., protecting identification of those who do not wish to participate), fairness, and methodology. Rounds (1996) reminds us that absolutes in ethical guidelines (i.e., those required by review boards) are “narrowly confined to getting as much information as possible from the researched while not exploiting or abusing them, and without violating their privacy or breaching confidentiality” (p. 53). But what about other issues that fall beyond these narrow absolutes? Those are the ones where individual decisions need to be made.

A more recent book (De Costa, 2016) is also based primarily on researchers’ experiences. The range of topics shows the change in emphasis in ethical concerns over the 20 years since Schachter and Gass’s edited volume, moving from procedural issues to issues of social justice and ethical behavior. For example, in De Costa’s volume, Starfield (2016) addresses the ethical issues involved in high-stakes proficiency testing; Bigelow and Pettitt (2016) present ethical dilemmas that arise when conducting research with immigrants who have limited formal schooling, and Thorne, Siekmann, and Charles (2016) confront ethical issues when conducting research with indigenous language populations. All of these areas go beyond the narrow confines of what is required by review boards, and the authors discuss contexts in which researchers are forced to make ethical decisions during a research project and beyond.

Responsible Conduct of Research

In addition to discussions of social responsibility of research, a second, and perhaps better known emphasis, has been placed on a particular sub-branch of research ethics, the *responsible conduct of research* (RCR). RCR education was first recognized in the 1989 Institute of Medicine report, and since then has evolved into a strand of research ethics for all disciplines using human participants (Steneck, 2007). With regard to Justice Stewart’s statement

provided at the beginning of this chapter, RCR is *what we have the right to do*, but taking into account the broader picture of social responsibility focuses on the *right thing to do*.

RCR was first established in the US during the 1980s in response to ethical scandals that had come to light in the previous decade (Steneck & Bulger, 2007). The rapid development of technological, social, and scientific advancements at the time (see Broad & Wade, 1983 for a then-current perspective on these issues) resulted in the promulgation of RCR guidelines, but it was not until the early 2000s that the Office of Research Integrity, part of the US Department of Health and Human Services, adopted a formal RCR training program (Steneck & Bulger, 2007). The development of RCR training occurred simultaneously with the growth and development of IRBs in the US university system, producing an atmosphere that treated research ethics as necessary red tape for conducting research (Van den Hoonaard, 2011). In fact, Haggerty (2004) refers to *ethics creep*, which he defines as “a dual process whereby the regulatory structure of the ethics bureaucracy is expanding outward, colonizing new groups, practices, and institutions, while at the same time intensifying the regulation of practices deemed to fall within its official ambit” (p. 394). Only in recent years has IRB approval become an expected staple of the social sciences research landscape in the US and with comparable review boards in many countries around the world.

RCR itself was developed as a method of instruction for research ethics (Steneck, 2007) but has since developed into a strand of research ethics unto itself. The Association for Practical and Professional Ethics (<http://appe.indiana.edu>) includes RCR as part of its RCR and Research Integrity Ethics special interest track. RCR itself consists of nine domains (see Table 32.1), which do not cover all possible ethical issues, but which do allow for a simplified means of discussing a broad topic. It should also be pointed out that the domains are not exclusive, with frequent overlap between and among topics. As an example, publication name order might be an issue of mentorship, collaborative science, or authorship/publication.

Much of the discussion of research ethics in practice is based on a few reported instances from the field (for examples see De Costa, 2014; Hobbs & Kubanyiova, 2008; Lee, 2011; Li, 2011; Shohamy, 2004). While these reports are invaluable, they do carry the bias that they are typically written only when ethical problems arise. Bowern (2010) sensed a similar issue in the reporting by linguists of interactions with local IRBs. However, when linguists were surveyed and asked for their general experiences with IRB applications, most responded with very little negative criticisms, the largest being that the use of children in research was difficult and the amount of time added to a project increased.

Table 32.1 RCR domains

RCR Domain	Issues included
Human protection	Consent, confidentiality, and risk/benefit ratio during data collection
Publication/authorship	Receiving just credit for work, citing others, duplication of research
Research misconduct	Falsification, fabrication and plagiarism of data
Animal resources	Humane treatment of lab animals
Mentorship	Training/supervision of graduate students, team management
Data management	Ownership, collection, safety, use, and sharing of data
Collaborative science	Sharing of results, data, and credit among research colleagues
Conflicts of interest	Financial, personal, intellectual stakes in project
Peer review	Accurate and timely review of submitted work

Susan Gass and Scott Sterling

As noted earlier, in L2 research, both social relevance and attention to procedural ethics are prominent. Sterling et al. (2016), in an effort to produce quantitative data on research ethics in the SLA context, investigated the type and amount of research ethics training received by applied linguists and SLA scholars. Using a survey approach, the authors found that materials related to IRB applications and training received high levels of attention in graduate programs and in professional training. These IRB issues related to four RCR domains (see Table 32.1): (1) data protection, (2) conflicts of interest, (3) human subject protection, and (4) research misconduct. The other domains received far less formal or informal focus. This trend was found to exist over time, meaning that even with an increase in the overall education in research ethics, far less time has been spent on non-IRB related issues during graduate training in SLA.

Other quantitative data-driven research into research ethics has focused on the need for (1) accurate and thorough reporting (Polio & Gass, 1997), and (2) the proper usage and reporting of statistics in SLA research (Loewen et al., 2014; Norris & Ortega, 2000; Plonsky & Gass, 2011). Proper reporting generally falls under the authorship and publications or the research misconduct domains of RCR, and journal editors are often the guards in such instances.¹ Whether intentional or not, the misuse of statistics can lead to the development and maintenance of scientifically unsound theory. Plonsky and Gass (2011) found multiple issues with the reporting and study quality of SLA studies, which is not surprising, given that Loewen et al. (2014) found that while many researchers had taken courses on statistics, many felt ill-prepared to actually use them.

Key Concepts

Procedural ethics: Ethical considerations undertaken for approval of research.

Research integrity: Ethical considerations applied to all aspects of research from planning to publishing. Often considered to include both professional duties and ethical responsibilities.

Research ethics: Ethical decisions encountered during any step of the research process.

Responsible conduct of research (RCR): A strand of research ethics that focuses on both procedural ethics and researcher integrity.

Ethical review board: A general term for an organization that verifies that research involving human subjects meets minimal governmental or institutional guidelines.

Institutional review board (IRB): A particular instance of an ethical review board found in the US. IRBs are generally housed inside of universities and are tasked with approving research studies if they are in compliance with federal, state, and institutional regulations.

Empirical Evidence

What Do We Know and Where Can We Go?

Often discussions of research ethics are in some sense sanitized with there being clarity as to what is and what is not ethical. This follows on work by Sterling et al. (2016), where ethical training emphasizes some aspects of research behavior and minimizes other

behaviors. Sterling and Gass (forthcoming) conducted a broad survey that investigated the ways in which researchers (faculty versus students; experienced researchers versus less experienced researchers) reacted to ethical issues in classroom-based research. Data for their study came from a survey in which 10 scenarios² were created (based on seven of the nine RCR domains: human subjects protection, publication/authorship, research misconduct, data management, collaborative science, conflicts of interest, and mentorship; and three scenarios involving consent issues that arise in classroom contexts: the language of the consent forms, consent forms used in non-USA contexts, and hiding research agendas from participants). Among other questions, respondents were asked to evaluate each situation based on a 6-point scale ranging from completely unethical to totally ethical. As a way of taking the ethical pulse of the respondents, we note two important factors that emerged from the Sterling and Gass (forthcoming) study, one concerning the scenarios themselves and one concerning the level of research experience of the respondent.

First, four scenarios (research misconduct, conflicts of interest, consent in a non-US context, and the language of consent) were viewed as unethical³ and four (publication/authorship, data management, mentorship, and hiding the research agenda from participants) were viewed as ethical. The other two were perceived as neither ethical nor unethical.

The division roughly correspond to two types of topics outlined by Guillemin and Gillam (2004): procedural ethics and research conduct (academic integrity). The former are more likely to be covered in IRB training and because of this, researchers may be sensitive to those issues. The latter, on the other hand, develop organically, with researchers typically receiving no specific training on how to address them. The results of Guillemin and Gillam (2004) correspond with those reported in Sterling et al. (2016), who found that formal training in graduate school emphasized areas of research misconduct and human subject protection with much less emphasis on mentorship, collaboration, authorship, and peer review. In other words, procedural ethics generally receive formalized attention whereas areas of professional conduct do not. In fact, participants in the survey conducted by Sterling and Gass often mentioned that scenarios involving non-IRB issues seemed unprofessional but not unethical in nature. We suggest that the relative lack of attention to academic integrity (mentorship, authorship) makes that area less obviously unethical to participants.

A second point of note is that in the Sterling and Gass study, in general, those with more classroom-based research experience viewed the scenarios as more ethical than those with less experience. Sterling et al. (2016, p. 31) point to a possible “connection between the rate of occurrence and the perceived ethicality of each scenario.” They also note that the respondents in their study commented on the fact that “particular scenarios were more ethical because they occur more often.” It is possible that as researchers experience more instances of issues that might fall in a so-called grey area, they become more conditioned to accept issues that they otherwise might not have accepted in earlier stages of their research careers. In other words, desensitization to what otherwise might be (borderline) unethical behavior occurs with frequent encounters.

We turn now to a consideration of grey areas that are important for those planning research projects. We posit eight questions for the reader to consider, and we then offer possible advice for ways of conceptualizing each issue. There are no solid answers to these

Susan Gass and Scott Sterling

questions, but instead researchers must critically think through these issues as they plan, conduct, report, and then end a research project.

1. *What to do if a student/students do not want to participate in a study?*

Issue

Classroom-based research by its nature is conducted in a classroom with multiple autonomous individuals. The agency provided to each of these individuals varies depending on their age, culture, and other factors. Do children have the right to consent (legally or morally) to a research project or can a parent do it for them? If the latter, do children have to assent as well? Can some students consent to being part of the project while others decline?

The final question is reflected by a specific comment from a respondent in the Sterling and Gass (forthcoming) study.

Students chose to take the class, but not the research. If the student wishes to fully participate in class, their voice is likely to show up in the video, which makes them a part of the research.

Considerations

Some possible recommendations are to move students off-camera or to delete any instance of their voice from a recording. But an even more difficult question to answer is what happens if a student declines participation in a research project and due to the research question, the teacher agrees to change his/her teaching style—are we then forcing the student to participate in research, even without his/her data being collected? This problem is exacerbated when the use of an experimental teaching method is introduced, especially if the results of using this method are unknown.

2. *What if understanding the research context for publication requires revealing too much information about participants?*

Issue

What should be the main concern of researchers? (1) reporting detailed data that allows for understanding of the research context and comparison of data across studies, or (2) the absolute protection of participants' confidentiality? For example, what about a case where a language instructor's behavior might reflect poorly on him/her or the place of employment? In some instances absolute confidentiality might not be possible.

Considerations

When describing the venue of data collection, it is often the case that it will become obvious to readers (or at least to other people who were present during data collection) where the data collection took place and perhaps even who the teacher was (if the teacher was the focus). A similar situation was reported in Lee (2011), where the author observed various instances of racialization occurring to her primary participant. Lee had to decide how to report the data without impacting her participant. To do that she opted to wait a number of years until her participant had changed jobs before she published the results.

3. *What if a researcher wants to use data for a different use than participants originally agreed to? Does participant consent continue for all secondary uses of data?*

Issue

Consent is a significant part of any research project. Ideally, if the data need to be used differently or if an expanded use of the data might provide important information or an important service to the research community, participants should be asked for further consent if possible.

Considerations

Asking for retroactive consent is often impossible, if not impractical, once a school term ends and students are dispersed to a new class or school. If data are collected in a high school, the researcher would only have 1–4 years before graduation to collect new consent before tracking down participants becomes almost impossible. But, what if technology becomes such (e.g., the establishment of corpora) that one can do something different (and useful): should the original participants be asked, especially if anonymity can be preserved? What are the boundaries between legitimate and nonlegitimate use of old data, or is this purely up to the researcher if IRB has no significant input? Applied linguistics and SLA researchers are not alone in grappling with this issue; similar questions are being asked in other fields, such as biobank data in genetic research.

4. *How can researchers balance issues of conflict of interest when conducting research in his/her own classroom? If an instructor is simultaneously teacher, course designer, and researcher, how does one ensure that student interests are paramount?*

Issue

When it comes to improving and understanding language pedagogy, action research and instructed SLA projects are critical to the field.

Considerations

In terms of research ethics, it is key that researchers establish a well-thought-out protocol that protects student and other stakeholder interest. Hopkins (2014) is unequivocal when he states “the teacher’s primary job is to teach, and any research method should not interfere with or disrupt the teaching commitment” (p. 70). Likely any research into a classroom will cause some disruption, from the presence of a videorecorder to a change in the materials to be covered. Hopkins further provides the following “checklist” of steps to follow when conducting action research. These guidelines are important in the planning stages before any actual research takes place (see also the papers in Schachter & Gass, 1996, where many of these same issues were addressed).

Guidelines for Good Classroom Research (Hopkins, 2014, pp. 72–73)

- Observe protocol.
- Involve participants.
- Negotiate with those affected.
- Report progress.
- Obtain explicit authorization before you observe.
- Obtain explicit authorization before you examine files, correspondence or other documentation.
- Negotiate descriptions of people’s work.
- Negotiate accounts of others’ points of view (e.g., in accounts of communication).

Susan Gass and Scott Sterling

- Obtain explicit authorization before using quotations.
- Negotiate reports for various levels of release.
- Accept responsibility for maintaining confidentiality.
- Retain the right to report your work.
- Make your principles of procedure binding and known.

It is important that any classroom intervention mitigate as many potential issues as possible before they even surface. Most important is researchers' acknowledgment that they are entering the classroom with their own opinions and biases and that these most likely have influenced the project and the questions being asked.

5. *How can mentor advice be evaluated?*

Issue

In the world of academia, education in research ethics is largely modeled by mentors (Steneck, 2007), a finding substantiated by Sterling et al. (2016), who found that much ethical training was conducted informally, likely by mentors. However, surveys (Sterling & Gass, forthcoming; Sterling et al., 2016) indicate that many SLA scholars are not well versed in research ethics and that mentorship is not a topic that is often taught to future mentors. Taken together, a situation arises in which largely untrained mentors are potentially passing on unevaluated ethical behaviors.

Considerations

Stories of misuses of mentorship in academia include advisees being pressured into situations, work being "stolen" or at least improper credit assigned, or poor advice that negatively affects the career or life of the mentee.

Mentorship is a key component for any researcher-in-training, especially when it comes to instructed SLA. Classrooms, instructors, and schools are a limited resource and so research in second language classrooms is often a one-time event. Because researchers-in-training do not have unlimited opportunities to collect data in a given timeframe, the project needs to be designed and executed in such a way that data are accurately captured. Reliance on the advice of a mentor is critical, but what if a student researcher does not feel comfortable with a mentor's advice, particularly as it relates to a classroom-based research project? One always has to be true to oneself and at the same time recognize that there are grey areas, especially when power dynamics are shifted in such a way that going against the advice of a mentor could have serious repercussions for the mentee. Open discussion is always the best strategy whether with one's mentor or with others who serve in a mentoring capacity.

6. *Consent forms in different cultural contexts.*

Issue

In conducting research, one often finds oneself in a position where cultural norms conflict. It is clear that ethical issues are often grounded in a particular cultural context (see Rounds, 1996; Swain & Cumming, 1989, who liken this to an anthropologist being a stranger in a strange land). Consent forms are not common in many countries and as a result, teachers, students, and administrators may be suspicious of the legalistic nature of consent documents. The request for participation from an instructor who is asking for students' voluntary participation may be interpreted as a requirement.

Issues of different interpretations of signing consent forms is particularly difficult for researchers based in the US, where IRBs have historically been known to be quite strict in the interpretation of procedural research ethics, and strongly affirm the belief that all participants should not only be made aware of research but should also be provided explicit instructions for joining, withdrawing, and all other rights thereof.

Considerations

While unlikely, it is possible that data obtained without consent will not be allowed to be used—meaning that researchers have misused the time and energy of participants. IRBs have been known to audit records to ensure that researchers are actually complying with regulations. One such instance of an audit might be instigated due to claims of misconduct or harm against a researcher.

Our advice for handling situations where signed consent cannot be obtained, or any IRB complication, is to use open communication. In many universities, the IRB is staffed by volunteer faculty members (often with expertise in law and/or ethics) who will likely understand that exceptions need to be made. When situations arise, the best course of action is to (1) triage the situation as best as possible (record verbal assent for example), and then (2) contact the IRB as soon as possible and ask for advice or possible exceptions to the rule. Waiting until a research project is finished before asking for assistance might compromise the ability to ask participants to consent in a different way, and might limit partially or entirely the use of the data collected.

7. *Consent forms for nonnative speakers.*

Issue

There is no uniform IRB requirement that consent forms be adapted for those who are not native speakers of the language of the form, although translations are sometimes required or suggested.

It is important that those who sign consent forms understand what is written therein. However, literature on the comprehension of consent forms largely shows that participants do not read, nor do they fully understand the documents they sign regardless of the language the form is read in. While this may be the case, it is important that consent forms have the potential of being understood even if an individual decides not to read them carefully.

Considerations

While translation is often a good option, it is not always the answer in that other problems surface, such as the need to translate forms into multiple languages, participants' potential inability to read in their native language, and/or a loss of meaning during translation. Sterling (2015) analyzed consent forms from many applied linguistics researchers who were conducting research with English-as-a-second-language (ESL) learners. He found that the majority of the consent forms analyzed were complex in terms of reading and vocabulary level and possibly beyond the ability of ESL learners to comprehend.

One cannot place the responsibility on IRBs for determining the comprehensibility of a consent form for nonnative speakers; it is actually quite unlikely that those approving IRB applications are language experts. And, writing a form in simplified language for native speakers is not the same as preparing a document for international students who might not share beliefs or background knowledge in research ethics. IRB members are most likely not attuned to the special needs of international or ESL students in the same way that researchers in the SLA community are. Asking IRB members to be expert language

researchers would be akin to asking an SLA researcher to judge the relative safety of using various sizes of beakers in chemistry. Providing comprehensible consent forms for participants will largely fall on the researcher, which, according to Sterling (2015) is not the current norm, at least in ESL research. Additionally, although it is not the area of interest for all SLA scholars, providing researchers in other disciplines information about the comprehensibility of consent forms for nonnative English speakers might be one area where SLA scholars can give back and help support people even outside of research.

8. *Consent forms that mask the research agenda.*

Issue

It is often the case that one cannot reveal everything about a research project for fear of altering behavior of the participants. For example, a research project looking at implicit vocabulary learning from reading might not want to be too obvious about the research goals for fear of learners focusing explicitly on unknown words.

Considerations

Often researchers will approach this issue by being vague rather than specific during the consent process. The tolerance for generalities may vary from researcher to researcher. For example, some might be comfortable saying something to the effect that they are looking at how people learn second languages whereas others might find that too misleading and say something such as understanding how second language learners read novel passages. In general, however, the description of the tasks (even if the goal is partially masked) must be sufficient for potential participants to make an informed decision about whether to participate or not. In general, questions of what being truly informed of research means is not an easy topic. Cameron, Frazer, Harvey, Rampton, and Richardson (1993) note that what constitutes an appropriate deception is at the discretion of the researcher and ethical approval boards—although an overreliance on the IRB to inform researchers of what constitutes reasonable deception is not advised. Many universities have policies in place for when deception is allowed and when the deception should be revealed to the participant, although the level of vagueness allowed in such statements is usually not regulated. Rounds (1996) reminds us, however, that the researched has no role in determining appropriate deception. Teachers whose classes are being researched can be deceived if the researcher believes it to be an insignificant, but necessary deception. The question of how ethical such a practice actually is depends on the context of the research, the level of deception, and the eventual disclosure of said deception.

Guidelines

Before embarking on a classroom research project, the planning process involves numerous decisions. Prior to data collection, decisions can, as De Costa (2014) illustrates, have large impacts on how smoothly and ethically the rest of the research project will go. While there are always unexpected twists in research, a well-designed plan can minimize many issues during and after data collection. The first question addresses issues of justification. Is this research worthwhile? Who will benefit? Who are the stakeholders and what ethical issues might be encountered? Once the ethical issues are identified, researchers need to carefully consider how they will address them. Taking the IRB application seriously is one way to start this process. Next, one needs to think critically of the questions being asked on the application, and consider the worst case scenario. Additionally, researchers need to try to imagine

how research will affect others. If sensitive data are kept on a laptop and that laptop is stolen, what will that mean? Often in school settings, many researchers are interested in soliciting participants from a common pool (e.g., a particular classroom). This can be problematic because this could ultimately affect the education of students. What will happen if a politician reads a research report based on classroom investigations? Could the results be used against the group being investigating (see Shohamy, 2004 as an example)?

More specific to empirical research are the requirements presented by Emanuel, Wendler, and Grady (2013) as a way of determining if research is ethical. Their first point expands on the need for social justification by including scientific value as an important criterion. Their emphasis is on procedural ethics, and they cover many of the criteria present in most IRB applications. The utility of research does not have to be a direct one-to-one correlation with pedagogy. For example, understanding how novice language learners (mis)use motion and path verbs in Spanish classrooms might be important for someone developing a hypothesis on vocabulary acquisition, even if the results do not seem overly useful for pedagogy. Both merit and value must be assessed by the individual researcher.

Seven Requirements for Determining if Research is Ethical (Emanuel et al., 2013)

1. Social or scientific value: does research have merit or will it lead to something useful?
2. Scientific validity: are research methods rigorous?
3. Fair subject selection: are participants selected fairly and not coerced to join the project?
4. Favorable risk-benefit ratio: are any harms outweighed by the gains made for science, society, or the individual participant?
5. Independent review: have results been examined and approved by knowledgeable members of the field?
6. Informed consent: are participants knowledgeable enough of the study to determine if they want to participate or not?
7. Respect for potential and enrolled subjects: are participants protected before, during, and after the study?

As we have pointed out throughout this chapter, ethical decisions go beyond mere IRB approval and extend to what is right in a broader sense. As researchers grapple with some of these issues, we present questions that one might want to consider when coming to grips with one's research. We are not suggesting a particular answer to any of these questions, for each context is different and each researcher comes to the research process with a different set of assumptions and experiences. What we are suggesting is that these questions guide one in thinking about the consequences of research conducted in classrooms.

Ethical Considerations for ISLA Studies Beyond IRB Approval

- Who decides if students can participate in research? Can students be in a class but not participate in the research?
- Does research intervention warrant disrupting class instruction?
- Can students distinguish the roles of researchers, teachers, or administrators?
- Does signing a consent form truly entail understanding of the research study?
- How much information should be disclosed to participants?
- How can results that could have a negative impact on the lives of participants be masked?
- What role do faculty play in the future development of new ISLA scholars?
- Data never speak for themselves. How will you frame the story that you *find* in the data?

Susan Gass and Scott Sterling

And finally, we need a specific way to deal with our own actions and the consequences of those actions. In thinking about this issue, the following questions might serve as a guide.

Are Your Actions Ethical? Consider These Questions as a Way to Start

1. How will researchers' actions affect the way they are perceived by other members of the research community?
2. As a researcher, how will you feel about your actions 6 months or 6 years from now?
3. Is there concern about what others might think if they found out you as a researcher made a particular decision?
4. When deciding on a course of action, a researcher must ask the question about the intentions of that course of action. Was it with the best intentions or was it motivated by other reasons such as stress, lack of time, or need for results?
5. Will actions, even if they are technically within the bounds of IRB approval, cause negative outcomes for others?

Conclusion and Future Directions

The sine qua non of any research project within or outside of the classroom is ethical behavior. We have seen the rise of review boards and published guidelines by professional organizations. The mandate of an IRB is to guarantee researcher compliance with regulated (often governmental) rules of research; it is not meant to be the moral compass that researchers rely on. Situations will always occur that are outside of the purview of the IRB or that are not noticed in initial applications. In the Sterling and Gass (forthcoming) study, it became apparent how pervasive the idea was that the IRB is the ethical guideline to be followed and if the researcher receives IRB permission, everything in the study is ethical. Unfortunately, this perception indicates a misunderstanding of the ethical role that each and every researcher is required to take.

In classroom-based research, there are numerous other issues to consider such as non-procedural ethical issues. Ensuring that research is conducted ethically and rigorously is important for a growing science like SLA, which will likely see an increase in the overall number of members in the field, as well as an increasing number of venues for publication and arenas for research. As research in classrooms continues to develop, it will be important to consider other non-IRB issues in training future researchers, including nonprocedural issues or controversies surrounding academic integrity (mentorship, authorship, collaboration, and peer review). Additionally, we hope that future scholars will continue the trend of collecting empirical evidence when formulating best practices in research ethics, instead of relying only on anecdotal evidence. We hope that we have raised the collective consciousness of the field through the issues presented in this chapter, thereby pointing the direction of future challenges and considerations necessary for ethically conducting research in second and foreign language classrooms.

Notes

1. In Loewen and Gass's (2009) timeline of statistical practices, they note editorial decisions regarding reporting and statistical rigor. There are now many journals that have taken on the "guardian" role as stated in their submission policy. For example, *TESOL Quarterly* 1992 (expanded in 2003) introduced a section titled Statistical Guidelines for authors to consider before submitting papers to the journal. Similarly, in 1993 *Studies in Second Language Acquisition* introduced a replication section

- in the journal to deal with issues of reliability and validity. In that same year, *Language Learning* in their Instructions for Contributors stated, “Manuscripts considered for publication will be reviewed for their presentation and analysis of new empirical data, expert use of appropriate research methods” (p. 151). In 1999 and 2000, *Language Learning* issued editorial statements of reporting.
2. The 10 scenarios are as follows:

Situation #1 [RCR domain: human subject protection]

Danny has planned a research project in which he will observe a classroom for an entire semester. In the first 4 weeks he will observe the instructor teach normally. During weeks 5–7, Danny will ask the instructor to read five research-based articles. Danny will then observe the class during weeks 8–12 to see if the articles had any impact on the instructor’s teaching.

Danny wants to video record each class period and plans to move students who do not wish to appear on camera off to the side of the class where no recording will take place. Danny has received IRB approval for this project but he is having doubts. He is not sure if it is OK to force students to participate in research. The teacher has agreed to take part in the research but the students have no say, beyond specifying that they do not want to appear on camera. The teacher is the focus of the study, but it will impact the quality of the instruction.

Situation #2 [RCR domain: publication/authorship]

Betty has been a part of research group that has been collecting data for several years from one high school district. They are ready to start publishing their findings. Betty’s subproject involves conducting follow-up interviews with students during their first year of university. Betty has found that the German program at the research site is not producing students who are ready for university-level German. At the site, there are only two German teachers and only one teaches upper-level German.

Publishing the results would mean that anyone involved in the project would know about the instructor’s poor teaching record and it might also damage the reputation and funding at the school where data collection is still occurring. Betty is conflicted between her academic duty to report findings while also safeguarding the participants/school. She is also unsure about what she might owe to future students who would have to go through the poor-performing German program. In the end Betty publishes the data while keeping the site as ambiguous as possible.

Situation #3 [RCR domain: research misconduct]

Sidney has posttest data following 7 weeks of a classroom treatment taken from over 250 students. She has run several statistical analyses but has not found results that she feels warrant inclusion into a paper. After she attempted to clean up the data, she found approximately 10 random cases that could be considered statistical outliers. Once those cases were removed her results looked more promising.

A second pass through the data showed Sidney that if she removed 15 additional cases, she would be able to obtain results she felt would be compelling for publication. The issue is that these 15 cases did not qualify as outliers; however, they all come from the same classroom. Sidney is considering removing all 27 cases from the classroom (the one with the 15 problematic cases) so that she can feel comfortable publishing her results. Her rationale is that there must have been something unusual going on in the classroom even though she could not determine precisely what that was.

Situation #4 [RCR domain: data management]

Sally has hours of video footage and transcripts from a project she conducted 5 years ago. After her program hired a new linguist who specializes in corpus linguistics, Sally has decided to turn her data into a searchable database. She has used the database to find examples of naturalistic speech to show during her classes but not for research.

Mary, one of Sally’s students, wants to use the database to search for instances of certain collocations. Sally agrees to let Mary use the database and Mary applies for and is granted IRB approval to use the database as already existing data.

Situation #5 [RCR domain: collaborative science]

Five months ago Beth, a private language school ESL teacher, had no idea what communicative language teaching was. William had emailed her and asked if she wanted to take part in a research project. Beth was excited about the prospect and agreed. William asked Beth to audit a TESOL class and keep a diary of her teaching while she learned about CLT.

Susan Gass and Scott Sterling

William then collected data in Beth's class and wrote an article using data from the diary as well as the quantitative data. William felt that Beth had gone above and beyond just being a participant in research and asked her to join the project as a collaborator. Beth had no training in research methodology but thought that a publication would look excellent on her annual report. William did most of the work and became annoyed when Beth asked questions or offered suggestions as to what she thought was happening in the data. In the end, William published a paper with both of their names but the manuscript was largely a product of William's thoughts and not Beth's.

Situation #6 [RCR domain: conflicts of interest]

Lara notices that her students are having trouble perceiving the difference between two adverbs. As both an instructor and a graduate student, Lara thinks this will be a great opportunity to collect data while she is teaching.

Because Lara is in charge of organizing her classroom schedule, she decided to accommodate her research by planning more grammar activities instead of listening and speaking activities. She also believes that she will be able to control the amount of language input the students hear by regulating her teaching and increasing the frequency of the targeted vocabulary, helping her research area at the expense of other linguistic features.

Situation #7 [RCR domain: mentorship]

Last night, Liliana received an email saying that the research site she had planned to use for her dissertation would not work out. Liliana had deadlines coming up for her dissertation proposal and was desperate to find a French classroom where she could collect data.

Liliana approached her advisor, Dr. Shields, who mentioned that she had a friend who directed French at a local high school. Dr. Shields was currently collecting data at the school and thought that the context would fit Liliana's research agenda. When Dr. Shields first started to collect data, the French director said that the program did not want many researchers to come and use precious classroom time. Dr. Shields advised Liliana to contact the French director and hoped that by using Dr. Shield's name, the French director would let Liliana collect data.

Situation #8 [RCR domain: human protection]

Ryan had been in Russia for over a month now and things were not going the way he had expected in collecting his dissertation data. Before he left the USA, he had received IRB approval to conduct the research. However, once he arrived in Russia he found that most instructors and parents were unwilling to sign a consent form and in fact grew suspicious when Ryan presented the consent document.

Ryan was told by the administration at his research site that informed consent forms were not needed; the administration's approval was enough to conduct the research. Apparently, people become suspicious about signing documents. Ryan could tell that everyone was genuinely interested in helping so he decided for cultural reasons that he would not require anyone to sign a form or give verbal consent to participate.

Situation #9 [RCR domain: human protection]

Jenny was collecting data from 10 different classrooms at an intensive English program. The vast majority of the students in the program were from Saudi Arabia or Japan, but in the 10 classes there were approximately eight different languages spoken. Even if Jenny had had the funding to translate documents into all eight languages, she would not have been able to find translators for two of them.

Jenny felt uncomfortable as she stood in front of a group of 13 ESL students who were struggling to read her informed consent form. In another minute or two she would ask if anyone had any questions about the document, but Jenny already knew that no one would ask anything and everyone would simply sign on the line. Jenny did make herself feel better by remembering that the consent form was based on Dr. Pen's form and that it had been approved by her local ethical review board.

Situation #10 [RCR domain: human protection]

Jill is studying the types and duration of off-topic conversations during in-class group work among tenth grade (15- to 16-year-old) Spanish L2 students. The students, parents, teachers, and administrators all signed informed consent forms at the beginning of the study.

Jill did not want to bias her study, so in the purpose section of the consent form she only stated that she was interested in the effects of group work during Spanish language lessons. Everyone knew that

Jill would be video recording the students and would make every effort possible to keep students' identities confidential. The whole research agenda was approved by her local ethics review board. Jill is excited by the amount and richness of the collected data. She found that the students were often off topic but was surprised at the explicit topics that the young students were discussing. She had no idea that students would be discussing such topics. Everyone had agreed to the project and signed the proper forms so Jill believes that she can publish the results.

3. A scenario was considered ethical/unethical if it fell above or below the midpoint of responses.

References

- Bigelow, M., & Pettitt, N. (2016). Narrative of ethical dilemmas in research with immigrants with limited formal schooling. In P. De Costa (Ed.), *Ethics in applied linguistics research: Language researcher narratives* (pp. 66–82). New York: Routledge.
- Bowern, C. (2010). Fieldwork and the IRB: A snapshot. *Language*, 86(4), 897–905.
- Broad, W., & Wade, N. (1983). *Betrayers of the truth: Fraud and deceit in the halls of science*. London: Century.
- Cameron, D., Frazer, E., Harvey, P., Rampton, B., & Richardson, K. (1993). Ethics, advocacy and empowerment: Issues of method in researching language. *Language & Communication*, 13, 81–94.
- Cumming, A. (2002). Assessing L2 writing: Alternative constructs and ethical dilemmas. *Assessing Writing*, 8, 73–83.
- Davies, A. (1997). Introduction: The limits of ethics in language testing. *Language Testing*, 14(3), 235–241.
- De Costa, P.I. (2014). Making ethical decisions in an ethnographic study. *TESOL Quarterly*, 48(2), 413–422.
- De Costa, P.I. (Ed.). (2016). *Ethics in applied linguistics research: Language researcher narratives*. New York: Routledge.
- Duff, P., & Early, M. (1996). Problematics of classroom research across sociopolitical contexts. In J. Schachter & S. Gass (Eds.), *Second language classroom research: Issues and opportunities* (pp. 1–30). Mahwah, NJ: Lawrence Erlbaum Associates.
- Economic and Social Research Council. (2015). ESRC framework for research ethics. Retrieved from <http://www.esrc.ac.uk/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2015/>
- Emanuel, E.J., Wendler, D., & Grady, C. (2013). What makes clinical research ethical? *Journal of the American Medical Association*, 283(20), 2701–2711.
- Fox, J., Artemeva, N., Darville, R., & Woods, D. (2006). Juggling through hoops: Implementing ethics policies in applied language studies. *Journal of Academic Ethics*, 4(1–4), 77–99.
- Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and “ethically important moments” in research. *Qualitative Inquiry*, 10(2), 261–280.
- Hafernik, J., Messerschmitt, D., & Vandrick, S. (2002). *Ethical issues for ESL faculty: Social justice in practice*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Haggerty, K. D. (2004). Ethics creep: Governing social science research in the name of ethics. *Qualitative Sociology*, 27(4), 391–414.
- Hamilton, L. C., Hartter, J., & Saito, K. (2015). Trust in scientists on climate change and vaccines. *SAGE Open*, 5(3), 1–13.
- Hamp-Lyons, L. (1998). Ethical test preparation practice: The case of the TOEFL. *TESOL Quarterly*, 32(2), 329–337.
- Hobbs, V., & Kubanyiova, M. (2008). The challenges of researching language teachers: What research manuals don't tell us. *Language Teaching Research*, 12(4), 495–513.
- Hopkins, D. (2014). *A teacher's guide to classroom research*. New York: Open University Press.
- Koulouriotis, J. (2011). Ethical considerations in conducting research with non-native speakers of English. *TESL Canada Journal*, 28, 1–15.

Susan Gass and Scott Sterling

- Kouritzin, S. (2011). Ethics in cross-cultural linguistic research. *TESL Canada Journal*, 28, 1–4.
- Lee, E. (2011). Ethical issues in addressing inequity in/through ESL research. *TESL Canada Journal*, 28(S15), 31–52.
- Li, Y. (2011). Translating interviews, translating lives: Ethical considerations in cross-language narrative inquiry. *TESL Canada Journal*, 28(31), 16–30.
- Loewen, S., & Gass, S. (2009). Research timeline: Statistical rigor in SLA. *Language Teaching*, 42(2), 181–196.
- Loewen, S., Lavolette, E., Spino, L.A., Papi, M., Schmidtke, J., Sterling, S., & Wolff, D. (2014). Statistical literacy among applied linguists and second language acquisition researchers. *TESOL Quarterly*, 48(2), 360–388.
- Magnan, S. (2005). From the editor: Introducing the special issue. *Modern Language Journal*, 89, 315–316.
- Ngo, B., Bigelow, M., & Lee, S.J. (2014). Introduction to the special issue: What does it mean to do ethical and engaged research with immigrant communities? *Diaspora, Indigenous, and Minority Education*, 8(1), 1–6.
- Norris, J.M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528.
- Ortega, L. (2005a). For what and for whom is our research? The ethical as transformative lens in instructed SLA. *Modern Language Journal*, 89, 427–443.
- Ortega, L. (2005b). Methodology, epistemology, and ethics in instructed SLA research: An introduction. *Modern Language Journal*, 3, 317–327.
- Ortega, L., & Zyzik, E. (2008). Online interactions and L2 learning: Some ethical challenges for L2 researchers. In S. Magnan (Ed.), *Mediating discourse online* (pp. 331–357). Philadelphia: John Benjamins.
- Plonsky, L., & Gass, S. (2011). Quantitative research methods, study quality, and outcomes: The case of interaction research. *Language Learning*, 61(2), 325–366.
- Polio, C., & Gass, S. (1997). Replication and reporting: A commentary. *Studies in Second Language Acquisition*, 19, 499–508.
- Rounds, P. (1996). The classroom-based researcher as fieldworker: Strangers in a strange land. In J. Schachter & S. Gass (Eds.), *Second language classroom research: Issues and opportunities* (pp. 45–59). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schachter, J., & Gass, S. (Eds.). (1996). *Second language classroom research: Issues and opportunities*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Shohamy, E. (2004). Reflections on research guidelines, categories, and responsibility. *TESOL Quarterly*, 38(4), 728–731.
- Starfield, S. (2016). Quotidian ethics in the neoliberal university: Research and practice collide. In P. De Costa (Ed.), *Ethics in applied linguistics research: Language researcher narratives* (pp. 53–65). New York: Routledge.
- Steneck, N. (2007). *ORI introduction to the responsible conduct of research*. Washington, DC: Government Printing Office. own.edu.
- Steneck, N., & Bulger, R. (2007). The history, purpose, and future of instruction in the responsible conduct of research. *Academic Medicine*, 82(9), 829–834.
- Sterling, S. (2015). *Informed consent forms in ESL research: Form difficulty and comprehension* (Unpublished doctoral dissertation). Michigan State University, East Lansing.
- Sterling, S., & Gass, S. (forthcoming). *Perspectives from the field on ethical behavior in applied linguistics and second language research*.
- Sterling, S., Winke, P., & Gass, S. (2016). Training in research ethics among applied linguistics and SLA researchers. In P. De Costa (Ed.), *Ethics in applied linguistics research: Language researcher narratives* (pp. 15–37). New York: Routledge.
- Swain, M., & Cumming, A. (1989). Beyond methodology: Behind research. In J. Esling (Ed.), *Multicultural education and policy: ESL in the 1990s* (pp. 88–106). Toronto: OISE Press.
- Tarone, E. (1980). TESOL research committee report. *TESOL Quarterly*, 14(3), 383–390.

- Thorne, S., Siekmann, S., & Charles, W. (2016). Ethical issues in indigenous language research and intervention. In P. De Costa (Ed.), *Ethics in applied linguistics research: Language researcher narratives* (pp. 142–160). New York: Routledge.
- Van den Hoonaard, W. C. (2011). *The seduction of ethics: Transforming the social sciences*. Toronto: University of Toronto Press.
- Wen, Q., & Gao, Y. (2007). Dual publication and academic inequality. *International Journal of Applied Linguistics*, 17, 221–225.
- Yeager-Woodhouse, D., & Sivell, J. (2006). Prepackaged tour versus personal journey: The meaning of informed consent in the context of the teacher-study group. *Journal of Academic Ethics*, 4(1–4), 189–203.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Index

- accentedness 249, 261–2, 264
 accuracy 18–20, 50–2, 54–6, 58–64, 101–2, 118–21, 134–5, 188–90, 303–5, 307–8
 action research 79, 274, 329, 548–9, 552–3, 556–7, 585
 activities: affective 172–3; referential 172–3
 age of acquisition (AOA) 468, 471–4
 analysis of speech unit 57
 anxiety: foreign language 341, 345–7, 433–5, 438, 441, 443, 460, 473, 547; foreign language classroom 435, 438
 approach: components 285–6; discourse analytic 527, 529, 531–2
 assessment: classroom-based 8, 181–5; cycle 187, 197; formative 181–2, 185, 197; school-based 327–8; summative 181–2, 185; task-based 327–8
 attention: explicit 196, 257, 347; initiating joint 566; joint 562, 566–9; responding to joint 566, 568
 attrition 379, 490
 audio-lingualism 261–2, 265, 281, 328, 330
 authenticity 109, 227–8, 233, 240, 363
 automaticity 17–18, 98, 207, 248–50, 253, 346
 automatization 16–20, 22, 25–6, 51, 191, 254, 348, 351–3

 Bardovi-Harlig, Kathleen 226–8, 231, 238
 behaviorism 261–2, 281
 bilingualism 88, 272, 469, 498
 biliteracy 100
 Bitchener, John 307–8
 bootstrapping 508, 517
 Bot, Kees de 249
 Bowles, Melissa 492, 495, 498–9
 breadth and depth of word knowledge 283–4, 290, 293–4

 CALF measures 304–5
 causation 514
 cognitive effort 33–5, 37, 39, 42–4
 cognitive fluency 246, 248–51, 253
 Collentine, Joe 340–2, 349, 352

 complexity 8, 19–20, 22, 41, 101, 161, 212, 216–18, 227, 236, 240, 253, 285, 293, 312, 345, 349, 370, 400, 437, 444, 451, 459, 470, 494, 508, 570; grammatical 52–3, 58, 61–2, 475; lexical 52–3, 61–3, 307, 475; task 50, 58–60, 64, 120–3, 129, 132–3, 137–9, 303–4, 308–9, 407–9, 505, 549
 Complexity Theory 249, 330, 455, 522, 533
 comprehensibility 126, 248–9, 262, 264, 267, 269–70, 587–8
 computer-mediated communication (CMC) 60, 129, 135–40, 235, 364–70, 564; synchronous (SCMC) 60, 129, 135–40, 564
 confidence interval 511–12, 514, 517
 consciousness 97, 148, 206–7, 211, 215, 233, 238, 476, 480, 590
 conversational facial displays 566

 debriefing upon return 345, 352–4
 De Costa, Peter 525, 528, 532–3
 degrees of errors 55
 DeKeyser, Robert 19, 23–4, 111, 212, 341
 delayed posttest 22, 25, 41, 44, 99, 101, 117, 137, 155, 158–9, 174, 189, 274, 475–6, 491–4, 499, 507, 514, 517, 550–1
 density 52–3, 63
 Derwing, Tracey 250–2, 257, 260–3
 design: dominance 526, 529; experimental 59, 476, 478, 506, 514; quasi-experimental 514
 Dewaele, Jean-Marc 346, 440–4, 446
 distribution on personality dimensions 438
 diversity 50–3, 57, 63, 103, 287, 309, 340, 470, 522, 536, 550
 Dörnyei, Zoltán 419, 421–9, 433–4, 554
 Duff, Patricia 386–7, 578
 Dynamic System Theory (DST) 61, 422–3, 425, 428–9

 ecological validity 26, 140, 188, 507, 514, 572
 effect size 22, 136, 209, 211, 214, 434–5, 440–4, 508–15, 517
 elicited imitation 11, 157–8, 562, 569–72
 Ellis, Nick 206

Index

- Ellis, Rod 109, 112–15, 206, 213, 215, 252, 307, 365–9, 473, 569
- emergentism 533
- emotional intelligence 345, 439, 442, 445
- engagement 10, 71, 78, 99, 137, 285, 288, 294, 325, 329, 364, 379, 387, 392, 425, 452, 548, 554
- English: for academic purposes 135, 305, 534; as a lingua franca 262–3, 323–4, 340, 342; as a medium of instruction 89
- epistemology 525–6, 529, 579–80
- Erlam, Rosemary 115, 213, 402–3
- ethics: procedural 578–9, 582–3, 589; research 524, 578–82, 585–7
- ethnicity 10, 232, 346, 381–5
- external standards 181, 185–7, 198
- external validity 22, 509, 514, 526
- face-to-face 77, 79, 129, 131, 135–6, 236, 238, 332, 364, 366, 379, 382, 456, 564, 567–8
- feedback: corrective 8–9, 20, 26, 72, 78, 93, 95–6, 114, 122, 127–33, 138–9, 182–4, 210, 268–71, 273, 311, 400, 402, 407–9, 469, 475–8, 480, 494, 496, 498, 516, 544, 546, 551, 555; explicit 72, 74, 132, 173, 236, 403, 407–8, 410, 459, 478, 492, 494; implicit 21, 72, 209, 236, 403, 407, 410, 477, 494; metalinguistic 129, 213, 403, 407, 478; written corrective 9, 187–9, 191, 194–7, 300, 302, 307, 509, 517
- fluency: perceived 247, 250–1, 255; utterance 246–50, 252, 256–7
- focus on form 5, 8, 112–15, 122, 128–9, 132, 136, 138, 166, 169, 173, 175, 209–11, 218, 227, 273, 330, 366, 410, 451, 459, 489–90, 495, 498, 506, 550
- Foreign Language Classroom Anxiety Scale (FLCAS) 435
- foreign language enjoyment 441, 446
- form-focused episode 136, 546
- Foster, Pauline 54, 252, 543
- Freed, Barbara 256, 340–1, 350
- Frost Multidimensional Perfectionism Scale (FMPS) 443
- Gardner, Robert 324, 420, 424, 434–6, 440, 476
- Gass, Susan 127, 139, 194, 547, 579–80, 583, 590
- global measures 54–5, 57
- Goo, Jaemyung 21, 117
- grammar translation 210, 280–1, 324, 328–9, 334
- group: comparison 23, 100, 270, 511, 514; control 24, 188, 226, 236, 302, 309, 312, 475–6, 479, 493–4, 507, 510–11, 514, 542–3, 549–51, 555; treatment 221, 226, 236, 270, 306, 507, 510–11, 514, 542, 551, 555
- Gurzynski-Weiss, Laura 131, 137, 453, 455, 457–61, 526
- Hawthorne Effect 552
- heritage speaker 488–9
- holistic scales 54–5
- Horwitz, Elaine 434–5, 437, 440, 443
- Housen, Alex 52–3, 55
- Hulstijn, Jan 37
- identity 92, 229, 260, 266, 299, 341, 345–7, 349, 369, 381–6, 391, 522, 532, 535, 547
- individual differences 21, 23, 27, 37, 61, 128, 137, 174, 176–7, 250, 305, 332–3, 341, 343, 345, 347, 369, 380–1, 396, 399, 405, 445, 451–6, 460–4, 479, 533, 549–50, 557; learner 452–6, 460–1, 463
- input: comprehensible 91, 127, 207, 260, 299, 347; enhancement 128, 211–13, 218, 233, 239–40, 371; flood 97, 171, 212–13
- institutional review board 11, 532, 543, 552–3, 577–8, 582
- instruction: content-based 87, 89, 496, 498; explicit 5, 22, 34, 111, 116–17, 205–6, 208–15, 217, 233, 236, 255, 260, 265, 270, 273–4, 281, 283, 289, 292, 326, 366–7, 400, 403–4, 408, 480, 490, 492, 505, 508, 516, 551, 587; formal 36, 342; form-focused 5, 98, 101, 117, 208, 210, 254, 270, 293, 324, 326, 469, 570; implicit 209–10, 213–14, 217–18, 236, 326, 334, 404, 508, 515; study abroad and English-medium 2, 9, 22–3, 27, 248, 256, 271, 308, 323, 332, 340–1, 344–6, 353, 379, 384–5, 388–90, 524, 530, 535
- instructor individual characteristics 10, 452–8, 460–3
- intake 21, 37, 43, 168–70, 173, 265, 371, 566
- integrativeness 419–20
- interaction: aptitude-treatment 25, 399, 410, 544, 550, 557; scripted 127, 563, 565
- Interaction Hypothesis 72, 111, 126–8, 269, 407, 454–5, 474
- interactive hand gestures 566
- intercultural awareness 342, 345–6, 350, 353
- intercultural competence 369
- intercultural development inventory 346
- intercultural sensitivity 341, 343, 346
- interdependence 53, 57, 371
- interlanguage analysis 231
- internalization 3, 148–51, 167, 301
- intersectionalities 381–2
- introspective measures 543, 545–6, 548
- Involvement Load Hypothesis 37, 41, 301–2, 306
- Kasper, Gabriele 226–8
- Kim, YouJin 132–4, 137, 306
- Kinginger, Celeste 385, 352
- knowledge: declarative 4, 16–19, 21–6, 111, 207, 209, 348, 407; explicit 4–5, 15–16, 22, 39, 41, 147, 152, 187, 191, 206–9,

- 213–15, 217, 303, 305, 348, 366, 401, 409, 474, 489, 491, 493, 499, 569–70; implicit 4–5, 16, 18–19, 21, 23, 187, 205–10, 212–15, 217, 366, 409, 474, 489, 491, 493, 569–71; metalinguistic 213, 401, 569–70; pragmalinguistic 224; procedural 8, 15–19, 22, 111, 251, 407; productive 284–5, 287; receptive 284–5; sociopragmatic 225
- Kormos, Judit 309, 408
- Krashen, Stephen 3, 5, 35–6, 126, 207, 260, 299
- Kuiken, Folkert 62, 189, 306–7
- L2 self: ideal 419, 422–3; ought-to 422
- language: aptitude 6, 10, 131, 174, 396–401, 403, 408–9; majority 89–90, 103, 488–9; minority 89, 488–9
- language related episodes 76, 132, 271, 489, 495, 541
- language teaching: communicative 3, 108–10, 128–9, 207, 281, 322, 348, 399, 591; communicative orientation of 536, 544, 557; content-based 8, 87–103, 229; task-based 3, 8, 88, 108–23, 129, 322, 325–8, 333–4, 451, 509, 551, 556; task-supported 108–10, 117, 326
- linguaging 72–3, 76–9, 151, 158
- Lantolf, James 71, 74, 78, 150–2, 531–2
- Laufer, Batia 287, 289
- learner autonomy 9, 260, 270–1, 273, 321, 328, 332, 353
- learning: explicit 15, 18, 20, 35–7, 40–2, 44, 175, 206, 403, 565; implicit 15, 18, 35–44, 110, 140, 205–6, 397; incidental 2, 17, 33–44, 110, 117, 288–90, 292, 403; intentional 8, 15, 33–6, 38–43, 288–9, 425
- Leow, Ronald 37–8, 43
- Li, Shaofeng 396, 400, 403, 407–8
- Lightbown, Patsy 24, 89, 469, 479, 482
- Loewen, Shawn 42, 126, 129–30, 136, 138, 213, 322, 333, 451
- Long, Michael 5, 72, 111–13, 126–7, 208, 210, 454, 474
- Lyster, Roy 20–2, 25, 95–6, 98, 100, 130, 270, 478, 480, 482
- McDonough, Kim 21, 131, 135, 555
- MacIntyre, Peter 251, 425, 434–6, 440
- Mackey, Alison 21, 117, 127, 131, 139, 407, 459, 475, 477–9, 542–3, 547, 555
- managing access to TL speakers 347
- maturational constraints 471–2
- measure of textual lexical diversity 57
- mediation 70–1, 73, 148–51, 158, 161, 192, 196
- memory: episodic 17; procedural 23; rote 396–8, 401–2; working 10, 43, 120, 128, 131, 137–8, 152, 159, 161, 174, 346, 355, 396–401, 404–12, 451–2, 456, 549, 550, 557, 571
- mental lexicon 172, 282–4
- mental representation 8, 166–8, 170–3, 175, 177, 564–5, 572
- metalinguistic awareness 93, 127, 476, 571
- metapragmatic information 227, 233–4, 237–8, 240
- minimalist 167
- mixed methods research 524, 528–9, 536
- modality 60, 112, 137, 152, 301, 306, 368, 494
- Modern Language Aptitude Test (MLAT) 174, 397
- modified output 25, 127–8, 130–1, 137, 194, 454–5, 477–8
- modular curriculum 116
- Montrul, Silvina 492
- motivation: amotivation 419, 426; demotivation 419, 422–3, 426, 428–9; extrinsic and intrinsic 419, 424, 445; task 419, 424–5, 428; teacher 419, 554
- Motivation Orientation in Language Teaching (MOLT) 423, 554
- multifaceted nature of language learning and teaching 523, 525
- multimodality 368
- Munro, Murray 250–2, 261
- Nassaji, Hossein 74
- neuroticism 438–41, 443
- Norris, John 22, 51, 53, 208–11, 516
- Noticing Hypothesis 117–18, 127, 207, 547
- null hypothesis significance testing (NHST) 508, 510, 514, 517
- observations 89, 91, 96, 100–1, 129, 147, 161–2, 184, 231, 261, 274, 281, 312, 326, 332, 348, 354–5, 390, 428, 470, 476–7, 481, 509, 525–6, 529, 532–6, 541–6, 549, 551–6, 568
- Oliver, Rhonda 474–80
- ontology 525–7, 529
- oral-for-oral 235, 238–9
- Ortega, Lourdes 22, 51, 53, 126, 146, 208–10, 214, 300, 579
- output 3, 9, 25, 63, 93, 111–14, 121–2, 127–32, 135, 137, 174, 194, 209, 211–12, 218, 235, 248, 254, 269, 290, 292–3, 299–300, 306, 311, 342, 344–5, 347, 364–8, 370, 418, 436, 454–5, 458–94, 474–81, 490, 493, 542, 550
- Output Hypothesis 127, 254
- pedagogical norm 9, 261–7, 272–3
- personality traits: higher order 438–41, 443; lower order 439, 442–4
- Philp, Jenefer 407, 475, 477
- picture description 17, 39, 60, 214, 252, 268, 404, 563, 565
- Plonsky, Luke 132, 508–9, 513, 516
- Poehner, Matthew 78, 146, 150–1, 191–2

Index

- Polio, Charlene 54, 61, 131, 303, 305, 308
 position: noninterface 5, 19, 207; strong interface 5, 207; weak interface 5, 206–7
 practice: autonomous 97–9; deliberate 17; guided 97–9
 pragmalinguistics 224, 227
 predeparture preparation 345, 352–4
 presentation-practice-production (PPP) 109–10, 113, 117, 120–1, 207, 211, 326
 pretest 50, 62, 75, 101, 127, 130, 153, 155, 158, 174, 211, 226, 231, 235, 255, 350, 473, 476, 479, 492–4, 506–7, 510–12, 514, 517, 551, 555
 proceduralization 17, 19–23, 25–6, 207, 348, 352
 processing: automatic 111, 489; controlled 206, 407; depth of 24–5, 33, 37, 39, 41, 43–4, 93; input 8, 126, 166, 168–71, 173, 176–7, 194, 551; principles of input 169, 173; transfer-appropriate 17, 24, 26
 prompts, tasks, and genres 20–1, 25, 55, 72, 75, 95–7, 130–1, 140, 190, 193, 195–6, 288, 300, 303–5, 307–8, 310–12, 368, 407, 478, 547, 563–5
 pronunciation 4, 9, 15, 21, 34, 131, 134, 226–7, 249, 260–75, 293, 329, 366, 456, 460, 462, 478, 495, 507–8, 514, 546, 555; model 262–3, 266–72, 272
 quasi-experimental studies 10, 50, 99, 130, 478, 542–3
 real operating conditions 111
 recast 20–1, 25–6, 72, 95–7, 128–32, 137–8, 170, 176, 210–11, 213, 227, 270, 299, 301, 400, 403, 407, 409, 459, 476–8, 494, 498, 549, 555–6, 567
 reflexivity 522, 525, 528, 532
 research integrity 581–2
 responsible conduct of research 580, 582
 Robinson, Peter 36, 58, 120–1, 132
 Saito, Kazuya 20–2, 267, 270, 446
 salience 213, 216, 268, 494, 498
 SA programmes design 354
 Sato, Masatoshi 20–1, 25, 62, 140, 479
 scaffolding 62, 71–8, 92–7, 187, 191–2, 310–12, 480
 Schema for the Orienting Basis of Action (SCOA) 150
 Schmidt, Richard 127, 170, 226, 265, 348, 418, 420, 547
 Schmitt, Robert 283–6, 290, 292
 Segalowitz, Norman 246, 249, 254–7, 346, 350, 352
 semantic formula 225, 231, 237
 semiotics 524–5
 sentence completion 563
 sentence recall 563
 sequences: chunks or 18, 111, 248, 250, 309, 365; formulaic 248, 250, 252–5, 292, 309, 405
 Sheen, Younghee 114, 116, 477, 556
 Shintani, Natsuko 109, 120–2, 211, 307, 366, 516
 silence or breakdown 5, 51, 55–6, 62, 72, 263, 267, 382, 571
 Skehan, Peter 55, 58–9, 90, 101–2, 112–13, 120–1, 252
 skill specificity 17, 24
 social class 154, 232, 381, 384–5, 387–9, 440
 social networks, individual networks 341, 344, 352, 365, 369, 380–1, 386, 391–2, 488, 530–1
 social turn 341, 380–1, 522, 535
 Sociocultural Theory (SCT) 8, 69–73, 75, 77–8, 146, 148, 192, 330, 522, 531–2, 535
 sophistication 52–3, 60, 63, 309–10, 364, 385, 388, 427
 Spada, Nina 22, 24, 208–10, 214, 218
 specific measures 54–5
 speech act 224–5, 227–32, 237, 239, 533
 speech event 225
 speed or rate 17, 19, 23, 26–7, 55–6, 60, 63, 248–9, 253, 293, 346, 348, 571–2
 statistical power 508–9, 514
 stimulated recall interviews 545–7, 555–6
 Storch, Neomy 75–6, 79, 133
 structural priming 11, 134–5, 562–6, 568–9, 572
 structure and agency 381, 383
 study abroad objectives 342
 Swain, Merrill 72, 75, 77, 90, 127, 151, 254, 299–300
 task-as-process 108
 task-as-workplan 108
 tasks: focused 99, 112, 117, 270, 274, 400; orienting 34–5; pedagogic 112–13; target 112–13; unfocused 112, 117; written discourse completion 235
 terminal (T) unit 57, 309
 thick description 527, 529, 531
 think aloud protocols 38–9, 44
 Tomita, Yasuyo 22, 24, 209–10, 214, 218
 trajectories 50, 63, 74, 181, 186, 194, 248–9, 251, 257, 309, 332, 380–1, 385, 387–8, 532
 triangulation 528–9
 Trofimovich, Pavel 135, 264, 267, 271
 Truscott, John 187–8, 286, 290, 302, 307
 type/token ratio (TTR) 57, 63
 ultimate attainment 469, 471, 473, 488
 Universal Grammar (UG) 3, 126, 167–8, 207, 281, 470
 Ushioda Ema 342, 419, 421, 426–7
 VanPatten, Bill 168, 170, 173–4, 176–7, 535, 551
 variation 22, 53, 60–1, 156–7, 161, 176, 208, 211, 218, 227–8, 232, 304–5, 310, 330,

- 332–3, 341, 343–5, 351, 437, 455, 473, 498, 533, 535, 571
- visuospatial sketchpad 159, 397–8
- vocabulary acquisition 280–1, 283, 285–7, 289–91, 293–4, 589
- vocabulary control movement (VCM) 281
- vocabulary knowledge scale (VKS) 40
- Vygotsky, Lev 71–2, 147–51, 155, 158, 161–2, 192
- willingness to communicate (WTC) 251, 256, 324, 341, 345–7, 369, 380, 437, 443–4
- word families 284–5, 291–2
- writing-to-learn-language 300, 310
- written-for-oral 235
- written-for-written 235
- zone of proximal development (ZPD) 70–1, 149, 192



Taylor & Francis eBooks

Helping you to choose the right eBooks for your Library

Add Routledge titles to your library's digital collection today. Taylor and Francis ebooks contains over 50,000 titles in the Humanities, Social Sciences, Behavioural Sciences, Built Environment and Law.

Choose from a range of subject packages or create your own!

Benefits for you

- » Free MARC records
- » COUNTER-compliant usage statistics
- » Flexible purchase and pricing options
- » All titles DRM-free.

REQUEST YOUR
FREE
INSTITUTIONAL
TRIAL TODAY

Free Trials Available

We offer free trials to qualifying academic, corporate and government customers.

Benefits for your user

- » Off-site, anytime access via Athens or referring URL
- » Print or copy pages or chapters
- » Full content search
- » Bookmark, highlight and annotate text
- » Access to thousands of pages of quality research at the click of a button.

eCollections – Choose from over 30 subject eCollections, including:

Archaeology	Language Learning
Architecture	Law
Asian Studies	Literature
Business & Management	Media & Communication
Classical Studies	Middle East Studies
Construction	Music
Creative & Media Arts	Philosophy
Criminology & Criminal Justice	Planning
Economics	Politics
Education	Psychology & Mental Health
Energy	Religion
Engineering	Security
English Language & Linguistics	Social Work
Environment & Sustainability	Sociology
Geography	Sport
Health Studies	Theatre & Performance
History	Tourism, Hospitality & Events

For more information, pricing enquiries or to order a free trial, please contact your local sales team:
www.tandfebooks.com/page/sales

 **Routledge**
Taylor & Francis Group

The home of
Routledge books

www.tandfebooks.com